Essays on Development Economics: Families, Child Human Capital, and Migration

by

María Gabriela Farfán Bertrán

Department of Economics
Duke University

Date: ____________________
Approved:

____
Duncan Thomas, Supervisor

____
Erica Field

____
Amar Hamoudi

____
Peter Arcidiacono

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor
of Philosophy in the Department of
Economics in the Graduate School
of Duke University

2014

ABSTRACT

Essays on Development Economics: Families, Child Human Capital, and Migration

by

María Gabriela Farfán Bertrán

Department of Economics
Duke University

Date:_______________________
Approved:

____
Duncan Thomas, Supervisor

____
Erica Field

____
Amar Hamoudi

____
Peter Arcidiacono

An abstract of a thesis submitted in partial
fulfillment of the requirements for the degree
of Doctor of Philosophy in the Department of
Economics in the Graduate School of
Duke University

2014

Abstract

This dissertation consists of 3 essays on development economics, with an overarching theme that relates to the economics of the family, child human capital, and migration. The three essays combine rigorous empirical strategies with the use of uniquely rich longitudinal data, the Mexican Family Life Survey, to advance our understanding of individual, household and family behavior.

Using these population-level data, the first chapter is an evaluation of a prominent anti-poverty program, Oportunidades, on child nutrition. Oportunidades was a leading intervention in targeting resources towards women and linking public transfers to investments in child human capital, and currently serves about one quarter of the Mexican population. To isolate the impact of the program, I draw on evidence from the nutrition and biology literatures regarding the biology of child growth, in combination with the timing of the roll-out of the program and the panel dimension of the data. Consistent with previous evidence, this analysis finds positive and sizable effects on children who live in rural communities incorporated at the beginning of the intervention. In contrast, the impact of the program in rural localities incorporated later in time and in suburban and urban communities are, at best, very modest.

The second chapter uses extensive information on non-co-resident family members, and variation in the spatial dispersion among them, to study the extent to

which Mexican families share resources across households and test different models of family behavior. I extend previous work by explicitly looking at families with different degrees of spatial dispersion among their members, *including* families with members spread across international borders. I adapt the collective model developed in the intra-household literature to model the family decision problem, and I analyze family behavior with respect to two sets of outcomes: household budget shares and child human capital indicators. The results suggest that the combination of looking at different degrees of spatial dispersion within families and different dimensions of family behavior is crucial to a precise understanding of inter-household decision-making.

The third chapter offers an in-depth description and analysis of the determinants of the incidence and magnitude of cross-border remittances by Mexican migrants living in the United States. While the investigation of international remittances has a long history in both the scientific and policy literatures, developing a full understanding of the motivations for and the impact of these transfers has been constrained by inadequate data. In this analysis I use recently-collected and extremely rich longitudinal data on migrants, and their families in Mexico, to predict transfers behavior. Results suggest that important differences exist between male and female transfers patterns, and that key variables related to the degree of connection of the migrant with Mexico (and the U.S.), such as the location of family members, expectations about returning to Mexico, or savings/assets holdings, are all important in explaining remittances patters.

As a member of the team that implemented the third wave of the Mexican Family Life Survey, this thesis is part of a broader collaborative research agenda with both colleagues and advisors. In particular, Chapter 1 is in collaboration with Maria Genoni, Graciela Teruel, Luis Rubalcava, and Duncan Thomas. The programming, analysis, and writing of this chapter, as well as any errors in this work, are my own.

To my parents

Contents

Abstract iv

List of Tables xii

List of Figures xv

Acknowledgements xvi

1. Oportunidades and its Impact on Child Nutrition 1

1.1 Introduction..1

1.2 Description of the program 7

1.2.1 Roll-out, eligibility, components, and take-up 7

1.2.2 Expected impact on nutritional status 12

1.3 Current evidence: findings and limitations 14

1.3.1 Oportunidades in rural areas..14

1.3.2 Oportunidades in urban areas...20

1.4 Data...23

1.5 Identification strategy 26

1.5.1 Community group 1: received the program between 1997 & 1998...................35

1.5.2 Community group 2: received the program between 1999 & 2002...................37

1.5.3 Community group 3: received the program between 2003 & 2005...................37

1.5.4 Discussion...38

1.6 Descriptive analysis 43

1.7 Results 49

1.7.1 Rural sector 50

1.7.2 Urban sector 57

1.8 Interpretation of the results 59

1.9 Conclusion 64

1.10 Supplement: analysis of selectivity in the measurement of height 68

1.10.1 Descriptive analysis 70

1.10.1.1 Missing height in 2002 70

1.10.1.2 Missing height in 2005 71

1.10.2 Regression analysis 73

1.10.2.1 Rural sector 73

1.10.2.2 Urban sector 77

1.11 Supplement: measurement error in locality date of entry 79

1.12 Primary figures 81

1.13 Primary tables 84

1.14 Supplementary figures 93

1.15 Supplementary tables 95

2. Extended Families across Mexico and the United States 107

2.1 Introduction 107

2.2 Background 115

2.3 Model 118

2.4 Data and descriptive analysis 125

2.4.1 Data structure 127

2.4.2 Family in the data 128
2.4.3 Interactions among non-co-resident family members 132
2.4.4 Outcomes of interest 136
2.4.4.1 Budget shares 137
2.4.4.2 Child human capital 139
2.5 Empirical specification 142
2.6 Results: household budget shares 146
2.6.1 Neighbor families: all households interviewed in the same locality in Mexico 149
2.6.2 Families spread across Mexico 151
2.6.3 International families: households spread across Mexico and the United States 152
2.7 Results: child human capital 154
2.7.1 Families in Mexico 154
2.7.2 International families 155
2.8 Robustness checks: child human capital 158
2.8.1 Additional controls 158
2.8.2 Alternative samples 161
2.8.3 Non-linear effects 161
2.8.4 Alternative outcomes (international families) 162
2.9 Conclusion 163
2.10 Primary tables 168
2.11 Supplementary tables 185

3. International Transfers by Mexican Migrants in the United States 203

3.1 Introduction 203

3.2 Data 208

3.3 Empirical specification 212

3.4 Descriptive analysis 214

3.5 Results 221

3.5.1 Results for males 223

3.5.2 Results for females 226

3.6 Conclusion 229

3.7 Primary figures 232

3.8 Primary tables 233

References 254

Biography 263

List of Tables

Table 1: Expected impact of Oportunidades on child height as a function of the cohort, locality of residence and year of measurement, by rural-urban sector 85

Table 2: Final sample used in the analysis - Children 1 to 3 and 5 to 7 years old, by rural-urban sector 86

Table 3: Oportunidades participation rates by locality groups 87

Table 4: Potential exposure to Oportunidades, by cohort and locality group 88

Table 5: Estimated impact of Oportunidades on child height, by type of locality and rural-urban sector 90

Table 6: Estimated impact of Oportunidades on child height - Rural sector 91

Table 7: Estimated impact of Oportunidades on child height - Urban sector 92

Table 8: Missing height in 2002 - Test of differences between the 2005 height distribution of children measured in 2002 and children not measured in 2002 96

Table 9: Missing height in 2005 – Test of differences between the 2002 height distribution of children measured in 2005 and children not measured in 2005 97

Table 10: Logistic regression (=1 if missing height) - Rural sector 98

Table 11: Logit model - Missing height for the sample of children used in the analysis – Odds ratio and Chi2 tests of key variables – Rural sector 100

Table 12: Logistic regression (=1 if missing height) - Urban sector 102

Table 13: Logit model - Missing height for the sample of children used in the analysis – Odds ratio and Chi2 tests of key variables – Urban sector 104

Table 14: Estimated impact of Oportunidades using different locality-group classifications 106

Table 15: MxFLS Structure 168

Table 16: Evidence of help provided across households within families 169

Table 17: Summary statistics of main variables used in the analysis, by family type..... 171

Table 18: Estimated effect of family resources on household budget shares 174

Table 19: Estimated effect of family resources on household budget shares, by family-type - Main specification 175

Table 20: Pareto efficiency test when analyzing household budget shares - Main specification 179

Table 21: Estimated effect of family resources on child human capital indicators, by family-type – Main specifiacion 181

Table 22: Pareto efficiency test when analyzing child human capital indicators - Main specification 182

Table 23: Estimated effect of family resources on child human capital indicators among international families, by location of the child and location of resources 183

Table 24: Estimated effect of family resources on child human capital indicators among international families - Alternative measures of resources... 184

Table 25: Household structure in Mexico and in the United States, by MxFLS round... 186

Table 26: Family structure, by family type 187

Table 27: Difference in mean family resources and outcomes of the analysis, by family type.. 189

Table 28: Remittances by adult migrants interviewed in the United States 190

Table 29: Number of pairwise tests statistically significant when analyzing household budget shares - Main specification 191

Table 30: Estimated effect of family resources on household budget shares, by family type - Linear specification.. 192

Table 31: Estimated effect of family resources on household budget shares, by family type - Non-linear specification 194

Table 32: Pareto efficiency test when analyzing household budget shares - Linear specification 196

Table 33: Pareto efficiency test when analyzing household budget shares - Non-linear specification 197

Table 34: Estimated effect of family resources on child human capital indicators - Additional covariates 198

Table 35: Estimated effect of family resources on child human capital indicators - Alternative samples 200

Table 36: Estimated effect of family resources on child human capital indicators - Non-linear specifications 201

Table 37: Estimated effect of family resources on child human capital indicators - Alternative outcomes 202

Table 38: Descriptive statistics of adult migrants interviewed in the United States, by gender 234

Table 39: Descriptive statistics of transfers sent by adult migrants interviewed in the United States, by gender 238

Table 40: Transfers sent to Mexico in the last year by male migrants 242

Table 41: Transfers sent to Mexico in the last year by female migrants 246

Table 42: Transfers sent to Mexico in the last year - Interaction between each covariate and a female dummy 250

List of Figures

Figure 1: Expansion of Oportunidades over time at the locality level 81

Figure 2: Graphical exposition of the identification strategy - Definition of cohorts and locality groups 82

Figure 3: Expected impact of Oportunidades on children's height in Group 1 communities 83

Figure 4: Description of locality groups using Census data 93

Figure 5: Proportion of children with measured height, by age and MxFLS round 94

Figure 6: Incidence of transfers sent to Mexico by Mexican migrants in the United States 232

Acknowledgements

First and foremost, I would like to express my immense gratitude to my advisor Duncan Thomas, for his constant guidance and support. He has been an exceptional mentor and an endless source of inspiration during my studies at Duke University.

I am also very grateful to each of my committee members - Erica Field, Amar Hamoudi, and Peter Arcidiacono - for their guidance and support throughout this process. I would also like to thank Alessandro Tarozzi for his continuous help, feedback and support, as well as V. Joseph Hotz and Elizabeth Frankenberg for their helpful comments and advice.

I owe a special thanks to the Mexican Family Life Survey team – Maria Genoni, Graciela Teruel, Luis Rubalcava, Duncan Thomas, and Andrea Velasquez - for their passion, commitment, and example during the collection of the third wave of the survey. Additionally, I would like to thank my friends and colleagues in the development lab – Maria Genoni, Andrea Velasquez, Verónica Montalva, Ryan Brown, Mike Burrows, Nick Ingwersen, Dan LaFave, Pete Katz, and Evan Peet - for their friendship, motivation, and the immeasurable amount of things that I learned from them.

I want to recognize the profound influence, help, and support I received from mentors and teachers during my undergraduate and graduate studies in Argentina. In particular, a special thank you to Mirta González, María Inés Lara, Ricardo Leiva, and

Virginia Vera at Universidad Nacional de Cuyo, and to Leonardo Gasparini and Walter Sosa-Escudero at Universidad Nacional de La Plata.

Thanks are also due to the Duke Economics Department, the Duke Graduate School, DuPri, and Hewlett Foundation for their financial support.

Finally, I am extremely grateful to my family, especially my parents, Norma and Jorge, and my siblings, Laura, Nicolás and Mariana, for their unconditional love and support. I am also extremely grateful to Merve for her invaluable friendship, inspiration and support. Lastly, I want to express a very special thanks to Jeremy for his untold patience, continuous encouragement, and endless support.

1. Oportunidades and its Impact on Child Nutrition

1.1 Introduction

Oportunidades (formerly PROGRESA) is a large-scale conditional cash transfer (CCT) program that was implemented in Mexico in 1997, and covered, by 2005, about one quarter of the Mexican population.[1] This on-going antipoverty program not only assists millions of families, but it also substantially contributes to household resources. It is estimated that program income accounts for about 25% of total expenditure in rural beneficiary households, and 15% to 20% of total expenditure in urban households.[2]

This intervention was the first of its kind. It links cash transfers with investments in human capital as a way to integrate both short-term and long-term anti-poverty policies. Cash transfers are given to satisfy current consumption needs, while investments in education, health and nutrition aim to establish the grounds for the program to have long-lasting effects. The program was originally implemented in poor communities in rural Mexico, and the implementation was coupled with a rigorous experimental design. Based on the positive effects found in original short-term evaluations, the program has massively expanded over the years, covering now virtually all the country. Additionally, this program started a new trend in the design of

---

[1] PROGRESA stands for *Programa Nacional de Educación, Salud y Alimentación* (Education, Health and Nutrition Program) and Oportunidades for *Programa de Desarrollo Humano Oportunidades* (Oportunidades Human Development Program).

[2] Rivera et al. (2004), Leroy et al. (2008).

poverty alleviation programs, with slightly modified versions now found across the globe.

In this chapter we evaluate the impact of the program on child nutrition. While this impact is manifested in the short-run, an improvement in the nutritional status of children has long-lasting effects. Evidence shows that malnutrition in early childhood is associated with deficits in cognitive development, greater risk of infant and child mortality and morbidity, as well as worse health status and lower earnings during adulthood (Martorell, 1999; Martorell et al., 2005; Strauss and Thomas, 1995).

Nutritional status could be measured in a number of ways. In this project, we choose height as our marker of interest for a number of reasons. On the one hand, there is a well-establish positive association between adult height and earnings. Controlling for a number of individual characteristics, taller individuals earn, on average, higher earnings, and this relationship has been found to hold across a wide variety of contexts. On the other hand, we know from the nutritional literature that height at age two is a very strong predictor of adult height. Also, it has been shown that height is sensitive to nutritional inputs during the first years of life but not later in life. As will be explained below, we will exploit this last characteristic in our identification strategy. Finally, height constitutes a long-term marker of nutritional status, as opposed to other indicators that measure short-term nutritional status. In summary, we have a marker that is responsive to nutritional interventions early in life (better nourished children are, on average, taller), a marker that measures long-term nutritional status (as opposed to short-term nutritional status), and a marker that has a well-defined welfare meaning (given by its link to adult earnings, among other outcomes).

There are a few papers in the literature that evaluate the impact of the program on child nutrition, and focus on height as one of their outcomes of interest.[3] However, all the available evidence is based on the Oportunidades evaluation data which only represents the poorest communities in Mexico. These data consist of two samples. A rural evaluation sample selected in 1997, at the time the program first started; and an urban evaluation sample, selected in 2001 at the time the program expanded to cover urban areas. Both baseline samples are complemented with a series of follow-up surveys. The expansion of the program followed an explicit geographic targeting rule, where more disadvantaged areas were covered earlier. Thus, a key aspect of these data is that it represents very specific geographic areas of Mexico. As a result, the current evidence only speaks to the impact of the program on communities incorporated early in time, but there isn't any evidence on the performance of the intervention in places incorporated later, which are relatively better-off.

We believe this is an important gap in the literature that needs to be filled, as the success documented on rural communities incorporated in 1997-98 stimulated the massive expansion that we witnessed over the last years. This is exactly what we do in

[3] See Parker et al. (2008) for a summary of the literature.

this chapter. In contrast to previous evidence, we use population-representative data from Mexico to look at this question. This allows us to perform, for the first time, an impact analysis at the national level, and assess whether the current evidence on specific communities extends to the rest of the country. In the rural sector, we provide with the first estimates of the impact of this program on children living in communities incorporated after 1998. With respect to rural communities incorporated at the beginning of the intervention, for which evidence exists, the timing of our data allows us to compute the estimated impact on a cohort of children younger than that used in the Oportunidades sample. In this way, we can see whether the original effects found in these communities remain over time. Something similar applies to the urban sector. We evaluate the program in communities incorporated in 2001-2002, as well as in communities incorporated later. In contrast to the rural sector, however, estimates for all localities in the urban sector are relevant contributions to the literature, as it is not clear that the Oportunidades urban evaluation sample is representative of the urban communities incorporated in 2002.

The data used in this study is the Mexican Family Life Survey (MxFLS), an on-going longitudinal survey that collects an extensive set of information on individuals, households and communities. We use baseline information collected in 2002, and the first follow-up carried out in 2005-2006. Important for our project, these data are, at baseline, representative at the national, rural-urban and regional level. Being height our main variable of interest, we note that anthropometric measures were taken by trained personnel from the National Institute of Public Health (INSP). We complement these household data with Oportunidades administrative records to identify in which year each of our MxFLS localities was incorporated into the program.

Just as with any non-experimental study, the main challenge in the analysis is to be able to isolate impact effects in the absence of experimental variation. To overcome this difficulty, we propose an identification strategy that is founded in biology. In particular, we combine insights from the biology of child growth, the timing of the roll-out of Oportunidades and the panel dimension of MxFLS. The nutritional literature has established that height is sensitive to nutritional inputs during the first years of life, but nutritional interventions have only modest effects on children's height after they reach a certain age. While the exact age cut-off is still controversial, it ranges between the ages of 2 and 4. Based on this evidence, and the fact that Oportunidades expanded over time, program exposure is defined as a function of the age of the child at the time Oportunidades was introduced to the locality of residence. The strategy basically consists of identifying cohorts of children that were exposed to the program and cohorts of children that were not and then performing an impact analysis at the community level. The panel dimension of MxFLS helps us dealing with the dynamics of height over the life course. By measuring children at two points in time, we can compare children of different cohorts but measured when they are at the same point in their life cycle.

This identification strategy is innovative within the Oportunidades literature, but the basic intuition behind it has been successfully implemented in the nutrition and economics literatures. The Oportunidades literature can be broadly divided into two branches. The rural evaluation sample started with an experimental design that divided communities into treatment and control groups, and the control group did not receive the program until 18 months later. As a result, there are experimental short-term impact effects for these rural places. In contrast, longer-term impact effects in the rural sector, as well as all estimates in the urban sector, rely on more traditional non-experimental methods. We will be able to partially assess the performance of our proposed identification strategy, by comparing our estimates for rural communities incorporated between 1997 and 1998 to those obtained in the experimental evaluation.

The remaining of the chapter is organized as follows. Section 1.2 provides a detailed description of the program and explains the channels through which Oportunidades is likely to improve children's nutritional status. Section 1.3 presents a short literature review that summarizes the main results and discusses the main limitations this literature faces. Section 1.4 presents the data used in the analysis. Section 1.5 describes the identification strategy adopted in this analysis. Section 1.6 shows some descriptive statistics that help interpret the results presented in Section 1.7. Section 1.8 discusses the results and interprets them in relation to the current evidence, and Section

1.9 concludes. Sections 1.10 and 1.11 present some supplementary analysis, followed by sections 1.12 through 1.15 with figures and tables.

1.2 Description of the program

1.2.1 Roll-out, eligibility, components, and take-up

Oportunidades started in 1997 in the poorest rural areas of Mexico and has gradually expanded to cover less marginal rural, and urban areas. By the end of 1999 the program covered approximately 2.6 million families in almost 50,000 localities, which represents about 40% of the rural population. By the end of 2002 the program was operating in 70,520 localities, in all 31 states of the country, reaching 4.24 million households. As of mid-2005, Oportunidades covered 5 million families - about one quarter of the Mexican population.[4]

The program links cash transfers with investments on different dimensions of human capital. The rational of such integration is that cash transfers help households improve their current poverty status while investments in human capital produce long lasting effects that would help break the intergenerational transmission of poverty. Additionally, transfers are targeted at women (whenever possible). The foundation of

[4] For information on the characteristics of the program, see Angelucci and Attanasio (2009), Behrman and Todd (1999), Gutierrez et al. (2003), Rivera et al. (2000), Skoufias (2005), Skoufias et al. (1999a), Skoufias et al. (1999b).

such policy lies on the idea that transfers made to women have a higher impact on children than transfers made to men.[5]

The intervention basically consists of three elements: a universal monetary transfer (food component), an educational component, and a health and nutritional component. The universal monetary transfer is a fixed amount of money that every beneficiary household receives, and aims to improve the food consumption and nutritional state of poor families.[6] The educational component consists of a pre-specified amount households receive for each child enrolled in grades 3 to 12.[7] The health and nutritional component consists of access to basic health care services to all household members, the provision of nutritional supplements to specific demographic groups and educational talks. The nutritional supplements are provided to pregnant and lactating women and children between 4 and 24 months. They are also provided to children between 2 and 4 years old if malnutrition symptoms are detected by clinic personnel. The educational talks are community meetings where trained nurses and physicians

---

[5] The evidence in the literature is not conclusive. However, Rubalcava et al. (2009) show evidence consistent with that hypothesis for the Oportunidades case.

[6] A new transfer called "Senior Adults" was added in 2006, which is given to each adult 70 years old or older. Starting in 2007, households also receive an "energy component", a cash transfer established to help them face expenses related to energy sources. Starting in 2008, the program added the component "Vivir Mejor" which constitutes a fixed lump-sum transfer to compensate for food-price increases. Finally, in 2010 the component "Infantil vivir major" was implemented. It constitutes a fixed transfer for each child 0 to 9 years old.

[7] In 2003 Oportunidades added the component "youth with opportunities". It consists of a savings account that can be cashed when students graduate from high-school (12 years of education) if they graduate before they turn 22 years old.

discuss topics related to health, hygiene, and nutrition issues and practices. All benefits but the educational transfer are conditioned on regular health check-ups for every household member, regular attendance to the educational meetings, growth monitoring of preschool children, and regular prenatal and post-pregnancy health care visits. The required periodicity of health check-ups and attendance to educational meetings varies by household member. The educational component is further conditioned on regular school attendance of school-age children.

There are two important differences between the rural and urban components of the program. The first difference has to do with the household selection process. As explained in detail below, enrollment in the rural sector was basically at no cost, while that was not the case in urban areas. The different selection processes resulted in almost universal take-up rates in the rural sector, but much lower take-up rates among eligible households in the urban sector. Keeping in mind these differences is very important, as they determine the characteristics of available control groups as well as the identifying assumptions needed to evaluate the intervention. The second important difference has to do with the evaluation design. The Oportunidades rural evaluation sample has a short-term experimental design component, while the urban evaluation sample does not.

At first, the program was only implemented in rural areas, defined as communities with less than 2,500 inhabitants. Among all rural places, eligible localities were selected based on a marginality index which was constructed with the information available in the 1990 Mexican Population Census and the 1995 Population Count (Conteo). In the localities deemed eligible Oportunidades carried out a census to collect information on every household. This information was then used to calculate a household poverty index and identify beneficiary households. Finally, eligible households were informed about their eligibility status. As enrollment in the program was pretty much automatic and at cero cost, 97% of eligible households were incorporated to the program.

Among a subgroup of the rural communities classified as eligible to receive the program in 1998, a group was randomly selected to receive the treatment right away while the other group was assigned to receive treatment 18 months later.[8] By the year 2000 both groups were already under treatment. As a result, a new control group was incorporated to the evaluation sample in 2003 in order to estimate medium term impact effects. This new sample of localities was selected based on matching locality-level characteristics. Finally, a follow up survey was implemented in 2007 to evaluate long term effects.[9]

---

[8] Treatment communities started receiving the transfers in May 1998 and control communities during late 1999 and early 2000.

[9] The complete list of surveys that are part of the rural evaluation sample is: ENCASEH survey in 1997 (data used to identify eligible households), ENCEL surveys every six months between 1997 and 2000, ENCEL follow-up in 2003, ENCEL follow-up in 2007.

In 2001, marginal urban areas were incorporated into Oportunidades and urban localities were incorporated from 2002 on. Similar to the case of rural communities, census data were used to identify eligible areas. However, due to the high degree of heterogeneity in socioeconomic status within urban localities, there are two important differences relative to the rural component of the program. With respect to the geographic targeting of the intervention, the geographic unit of interest was the Primary Sampling Unit (AGEB in Mexico) rather than the locality as a whole.[10] Only localities with at least one PSU with high concentration of poor households were selected, and the program was implemented in those places. Additionally, a different household selection process was implemented.[11] Instead of collecting information on every household, the program established registration offices in eligible areas and advertised the program through campaigns. Households interested in the program had to go to the registration offices on specific dates and answer an inclusion questionnaire. With that information households were immediately classified as qualified for the program or not. If qualified, they had to answer a second questionnaire and were visited later in their dwellings to confirm their eligibility status. As a result of this longer and costly process,

[10] In urban areas, an AGEB consists of a group of 1 to 50 blocks.

[11] Marginal urban areas incorporated in 2001 were still under the previous system and so the selection process was the same as that in rural areas. Similarly, the household selection process applied to rural areas incorporated to Oportunidades in 2002 or later was a variant of that applied in urban areas (Gutierrez et al., 2003).

take-up rates were much lower than in rural areas: administrative data suggests that initially, only about 50% of eligible households registered for the program.[12]

In contrast to the rural sample, the urban evaluation design is not experimental. A sample of poor blocks was selected in 2001 from the areas assigned to receive the program in 2002. The control group was selected based on a matching process from localities planned to be incorporated to the program in 2004.[13]

1.2.2 Expected impact on nutritional status

Clearly, one of the components of the intervention is specifically designed to improve children's nutritional status. Both pregnant and lactating women as well as young children are given nutritional supplements on a regular basis. Additionally, two other components of the program are likely to affect the nutrition of young children. First, the cash transfer which is expected to improve the quality of the diet families consume. Second, the educational talks where health and nutrition related issues and practices are discussed. Trained personnel on the subject lead these meetings and it is mostly women who attend them, two factors that are expected to maximize the impact on children's nutritional status.[14]

---

[12] We are not aware of any updated estimates of take-up rates in the urban sector.

[13] The complete list of surveys that are part of the urban evaluation sample is: ENCERLUB survey 2002 (baseline), ENCERLUB follow-ups in 2003 and 2004.

[14] Better access to preventive and curative health care services, pre-birth health check-ups and growth monitoring, may also improve nutritional status.

However, participation in the program does not necessarily result in better nutrition. To begin with, nutritional supplements need to be properly consumed. There is some evidence suggesting that in both rural and urban areas access and consumption was not universal. With respect to the former case, Behrman and Hoddinott (2005) report that, during the first years of the program, the percentage of children aged 4 to 48 months that had access to the supplements varies from 52% to 63%. Similarly, during the first years of the urban program, Neufeld et al. (2004a) show that about half of the children aged 6-23 months took the nutritional supplements at least once a week, and only about a quarter of lactating women did. With respect to the other two channels, their influence depend on the degree to which money is actually used to improve the nutritional quality of the food consumed and the extent to which women implement what they learn in the educational sessions.

Since the benefit structure of the intervention is quite complex, the purpose of this research, as well as all the papers summarized in the following section, is to see whether the program improves the nutritional status of children. We do not aim isolating the effect of any particular component of the program, but see whether overall, there is a positive impact.

1.3 Current evidence: findings and limitations

Now we summarize the existing literature that focused on child nutrition, measured by the impact of the program on height.[15] Given the big differences between the rural and urban components of the program in terms of timing, selection process and evaluation design, each part is analyzed separately.

1.3.1 Oportunidades in rural areas

Two studies evaluate the impact of Oportunidades after one year of exposure in rural areas. By exploiting the experimental design of the survey, Gertler (2004) analyzes the impact on children aged 12 to 36 months in 1999. He finds that children in treatment villages are 1 cm taller than children in control villages, but also finds no significant effect on the probability of being stunted (more than two standard deviations below the reference median). On the contrary, Behrman and Hoddinott (2005) find no significant impact of Oportunidades on child nutrition when using an estimation strategy that relies on the random allocation of the program. However, based on the evidence that a shortage of supplements did not allow every eligible children in treatment areas to have access to them (and the fact that some children with severe malnutrition symptoms in control areas did receive them), they estimate next a treatment-on-the-treated effect. They control for the fact that access to the supplements was not random among eligible

---

[15] Other health outcomes evaluated in the literature include: obesity, anemia, weight-for-height, BMI-for-age, birth weight, probability of illness.

children using child fixed effects estimators and find that the program did increase growth per year by over 1 cm on children 12-36 months. They also evaluate the effect on the probability of stunting and find that children who receive the supplements have a predicted probability of stunting of one-third that of control children.

Rivera et al. (2004) also estimate the impact of one additional year of exposure to the program. However, they compare children with two years of treatment relative to children with one year (instead of one year relative to no exposure as was the case of the previous two studies). They only find a positive impact on children 6 months or younger at baseline that live in the poorest households. These children are on average 1 cm taller than children the same age with only one year of exposure.

Neufeld et al. (2004b) incorporate the 2003 round into the analysis. Using matching estimates, the study compares children in both early and late intervention communities (those that started receiving treatment in 1998 and those that were incorporated to the program 18 months later) with children residing in the new control communities. The authors find that children 24 to 71 months old in 2003 in the former group grew 0.67 cm more on average than control children and the prevalence of stunting is 12.4% lower (both effects statistically significant). They also compare the effects of differential exposure using the original treatment and control groups. Children 48 to 71 months in 2003 were fully exposed to the program if born in early intervention communities, but only partially exposed if born in late intervention communities. An evaluation of this differential exposure reveals no significant difference in height-for-age or prevalence of stunting between these two groups.

Finally, Fernald et al. (2009) evaluate the effect of additional 18 months of exposure almost 10 years after original treatment communities started receiving the benefits. They use height measured in 2007 and restrict the sample to those children born between March 1997 and October 1998. Children in early intervention communities were around 1 year or younger when they started receiving the supplements whereas children in late intervention communities were more than 1 year old. No effects were found on height-for-age z-scores for the whole group, but there was an effect of about 1.5 cm on height in younger children whose mothers had no formal education.

This summary reveals a mixture of positive and null impacts which depend on the methodology used, the difference in the degree of exposure, the time at which nutritional indicators are measured and the subgroup of children evaluated. In order to understand these differences the most important caveats that the rural evaluation sample faces are presented next. Some of them are common to any study that uses the rural evaluation sample while others are specific to evaluations that use nutrition indicators.

Short-term impact evaluations exploit the randomized design of the program, which help control for unobserved factors that differ between treatment and control individuals. However, the randomization was done at the locality level whereas impact estimates are performed at the household or individual level. While treatment and control groups look alike at the locality level, Behrman and Todd (1999) evaluate the differences in pre-program characteristics between treatment and control households and find that there are small but significant differences between these two groups. Additionally, a recent study shows that attrition, which was mainly ignored in this part of the literature, could potentially affect the results. Teruel and Rubalcava (2007) show that treatment households are more likely to leave the sample by the year 2000 than control households. The authors re-estimate the impact on high-school enrollment presented in Schultz (2004) and find that correcting for attrition results in higher impact estimates. As a result, short-term impact estimates could be biased due to deviation from perfect randomization and/or differential attrition rates between control and treatment groups.

Medium-term impact estimates face additional concerns. In the first place, they rely on matching estimators that assume that the relevant differences between control and treatment individuals can be controlled for by including observable characteristics. One immediate threat to this assumption results from the fact that the new control group in 2003 resides in localities placed in different geographic areas than the original control and treatment groups. As a result, any region-specific factor that cannot be controlled for can bias the results. There is also evidence of significant differences in terms of 1997 characteristics between the original evaluation group (treatment and control) and new control group. Parker et al. (2008) show that such differences include demographic characteristics, dwelling characteristics, ownership of durable goods, and household head and spouse characteristics. This situation can be partially overcome using difference-in-differences matching estimators which help control for time-invariant unobserved characteristics. However, these estimators face an additional problem. The new sample was drawn in 2003, so pre-intervention information on the new control group is based on questions that ask this group about their situation in 1997. As a result, the possibility of recall bias should be considered as the difference-in-differences matching estimators rely on retrospective information. Finally, attrition rates are not low: 83% of the households are in both the 1997 and 2003 surveys (78% of the individuals)[16]. To the extent that people that remained in the sample are different than people that left in unobserved dimensions that are correlated with the outcome of interest, this is an important issue to keep in mind. This is very likely considering that the survey was not designed to follow households but to come back to original dwellings, which explains why more than 80% of the attrition between 1997 and 2003 can be attributed to changes of residence or migration (Teruel and Rubalcava, 2007).

Fernald et al. (2009) is the only study that uses the 2007 round. As the authors mention, the main limitation of the study are the high attrition rates. They found no

[16] Only 60% of the households (47% of the individuals) report information in every survey between 1997 and 2003.

differences in characteristics measured at baseline between treatment and control groups for the sample found in 2007, but there were some differences between household characteristics of children used in the analysis and those lost.

In addition to the caveats already mentioned, there are two things that are specific to the nutritional data used to assess the program impact on height.

In the first place, indicators of nutritional status were not collected as part of the general evaluation survey. The data was collected at different times and by different teams, under the supervision of the National Institute of Public Health. This seems to have introduced some complications when trying to link nutritional indicators with the rest of the household and individual information. Furthermore, because of the difference in the timing, the first available indicators are measured at a time some households have already received some transfers, and so they do not correspond to pre-treatment data.

Secondly, there is evidence of shortage in the availability of supplements in the first years of the program. Adato et al. (2000) report that the distribution and intake of nutritional supplements seem to have been the most serious operational problem of the health component of Oportunidades. In response to this, health personnel exercised some discretion in the distribution of supplements by especially targeting those children that presented the most severe malnutrition symptoms. As a result, access to this component of Oportunidades among beneficiaries was, not only not universal, but also selective.[17] This implies that short-term impact estimates and estimates of differential exposure between the original treatment and control groups estimate intent-to-treat effects and may explain the lack of significant impact in some cases. Behrman and Hoddinott (2005) provide some evidence of this.

1.3.2 Oportunidades in urban areas

The two main characteristics of the urban evaluation sample were already mentioned in Section 1.2.1: it does not follow an experimental design, and take-up rates were initially very low (around 50%). The challenges of dealing with a non-experimental sample were exposed in the previous subsection when explaining the limitations associated with the new 2003 control group. The second element, however, introduces a new challenge to the estimations because eligible households that decided to enroll in the program are not expected to be a random sample of the set of eligible households in urban localities. Angelucci and Attanasio (2006) argue that traditional matching estimators, designed to control for non-random assignment to the program, may give biased estimators in the presence of non-random participation. The reason is that matching estimators rely on the assumption that variables that determine both participation and outcomes are observed. They propose an IV-type estimator that takes both nonrandom assignment and nonrandom participation into account and apply it to

[17] Behrman and Hoddinott (2005) find evidence of selective access to the nutritional supplements, but no evidence of selective access to the other components of the program.

the case of food consumption. They find that the estimated impact changes significantly when they use a traditional matching estimator compared to their preferred estimator. Parker et al. (2006) also use a combination of matching, differences and IV to estimate the impact of Oportunidades on schooling, and show that impact estimates vary as a function of the comparison group chosen. There is, however, no evidence on how much this would affect impact estimates on health outcomes.

Similar to the rural case, the use of nutritional data introduces additional concerns. Children in the data with nutritional information were not randomly chosen from the urban evaluation sample, but explicitly chosen to minimize the number of geographic areas in order to save costs. As a result, control children are not children who live in communities not yet incorporated to Oportunidades at that time, but children from eligible households that reside in the same communities as treated children but who did not enroll in the program. Even though this eliminates any bias due to locality-specific effects, it significantly raises concerns related to self-selection bias, especially considering what was mentioned in the previous paragraph. Additionally, height was measured both in 2002 and 2004 only on children younger than 2 years old at baseline, which limits the possibility of evaluating the effect on different age groups or performing robustness checks - the last of which seem to be crucial given the limitations of the data just exposed.[18]

Based on this data, Leroy et al. (2008) evaluate the impact of Oportunidades in urban areas on children younger than 24 months at baseline, in 2002. They use a two-year panel of 432 children and implement a difference-in-differences propensity score matching estimator. After two years of program exposure Oportunidades seems to have had no impact on growth in children 6 to 24 months but a positive impact on children less than 6 months old: the height-for-age z-score of the latter group is 0.41 higher than that of control children. They claim that selection bias is not likely to affect the results given that no significant differences were found at baseline between control and treatment groups in terms of height for children 2 to 4 years old. They also claim that loss of follow-up, which was 45% and 40% for control and treatment children respectively, is not a concern because there are no significant differences in baseline characteristics between children lost and children used in the analysis. Even though these robustness checks are encouraging, the importance and magnitude of the issues call for reservation.

---

[18] Children's mothers were also measured. Children 2 to 4 years old at baseline were measured in 2002 but not in 2004.

1.4 Data

The main data source of this research is the Mexican Family Life Survey (MxFLS). MxFLS is an on-going longitudinal survey that collects a rich set of information on demographic and socioeconomic characteristics of individuals, households, and communities. At baseline, the sample is representative at the national, rural-urban and regional level. The first wave (MxFLS1) was conducted in 2002 and interviewed 35,677 individuals in 8,440 households. These households reside in a total of 150 communities located across 16 different states.[19] The second wave (MxFLS2) was conducted during 2005-2006 and achieved a 90% re-contact rate at the household level. This wave consists of 36,946 individuals and 8,434 households, who due to migration decisions are located across 247 localities in 21 states throughout Mexico. The third wave (MxFLS3) started in 2009 and is now in the final stages of the field work. This analysis only uses the first two rounds of the survey.

The main variable of interest in this project is height, as it constitutes a marker of early life nutritional investments. MxFLS has anthropometric measures (weight and height) for every household member, and these measures were taken by trained personnel from the National Institute of Public Health (INSP). In order to control for age-gender specific differences, height-for-age z-scores are constructed using the 2000 CDC Growth Charts for the United States provided by the National Center for Health

[19] Mexico is divided into 31 states and the Federal District.

Statistics (NCHS). To that end, age in months is calculated using date of birth and interview date.

The survey collects information on Oportunidades participation at the individual, household and community level. We will exploit this self-reported participation but only in part of the analysis, as it cannot be used in our preferred specification. In contrast, most of our analysis follows an intent-to-treat approach. By drawing on evidence from the nutritional literature, program exposure will only depend on the age of the child at the time Oportunidades arrived to the place of residence. In order to identify the year in which each MxFLS community was incorporated to the Oportunidades program this analysis combines MxFLS data with Oportunidades administrative records.[20]

To this purpose, we obtained two administrative sources of information. One is the complete list of Oportunidades beneficiaries (Oportunidades' padron) as of December 2009. These records have individual information on locality of residence and date of enrollment in the program. Based on the households' date of entry, each of the 246 MxFLS localities is associated with the year in which the largest number of households was enrolled in the program. We will refer to this source of information as *Mode Year* data. An alternative source of information is a locality-level data with the

---

[20] The MxFLS community questionnaire has information on whether the community has or not the program at the time of the interview, but no information on the year of incorporation.

number of families enrolled each year.[21] We will refer to this source of information as *Expansion* data. Using both data sets, together with the self-reported participation rates recorded in MxFLS, each locality is assigned a year of incorporation.[22] Figure 1 illustrates the pattern of expansion of the program for the rural and urban areas of the country separately.[23]

To classify communities as rural or urban we use the 2000 Mexican Population Census. Following the official definition, which is the same Oportunidades adopts, rural communities are defined to be those with 2,500 inhabitants or less.

Finally, robustness checks will complement the definition of program exposure based on age and place of residence with other household characteristics. As mentioned before, one such variable will be self-reported participation as measured in our data. An alternative variable will be the eligibility criteria that Oportunidades follows to select beneficiary households, information that we were able to get through contact with Oportunidades administrators. The household poverty score is a function of household head characteristics (such as education and gender), dwelling characteristics (such as

---

[21] Due to some data entry error this source has missing information for the years 2002 and 2003.

[22] There is clearly some measurement error in this variable. 74% of the localities are assigned the same year regardless of the data source, 86% in the rural sector and 62% in the urban sector. In Supplementary Section 1.11 we show how the main conclusions are insensitive to using alternative years of incorporation.

[23] Officially, Oportunidades expanded to urban areas in 2001. The vast majority of households in urban areas that enrolled in the program before 2001 are either in semi-urban areas (2500-5000 inhabitants) or areas classified as rural before the 2000 Population Census (Parker et al., 2008). The characteristics of the urban localities that are assigned a year of incorporation before 2001 are consistent with this evidence, except for two localities, that have around 10,000 inhabitants.

floor material, access to water and sanitation), and other characteristics, such as overcrowding, possession of certain durables, or access to social security benefits. Fortunately, for the most part, the household-level score is not a function of variables that are very sensitive to the instrument used to collect the information, so that measurement is likely to vary substantially across different data sources. This would be the case, for example, if the poverty score was a function household income or consumption. Additionally, most of the household characteristics that go into the index are relatively stable over time. Both of these attributes make the score less sensitive to the data used to compute it as well as to the time at which is computed, which adds some validity to the implementation of the Oportunidades criteria to MxFLS data.[24]

1.5 Identification strategy

The identification strategy we propose in this study exploits the combination of three elements: the evidence that nutritional interventions have only modest effects on child height after the first few years of life, the fact that Oportunidades was not introduced in every place at the same time but gradually expanded over the years, and the panel dimension of MxFLS. Based on the first two, program exposure is defined as a function of the age of the child at the time Oportunidades was introduced to the locality

[24] Admittedly, there will still be some measurement error in the computation of this variable. However, we believe this is not a major concern for this analysis, as we only use the eligibility score as one of the criteria to select children from more disadvantaged households or households more likely to benefit from the intervention. The identification strategy, however, does not rely on comparing children in eligible households with children in non-eligible households.

of residence. The idea basically consists of identifying cohorts of children that were exposed to the program and cohorts of children that were not, and then performing an impact analysis at the community level. Variations of this empirical methodology have been successfully implemented in the nutrition and economics literatures (see, for example, Martorell and Habicht, 1986; Duflo, 2001; Frankenberg et al., 2005).

The biology of child growth defines a critical period over which nutritional interventions can significantly affect child height. While height is very responsive to nutritional inputs during the first years of life, height is pretty much determined after a given age. There is not absolute consensus in the literature regarding the exact age cut-off over which height is predetermined, but it does lie within the first years of life. In particular, the age range under discussion is between two and four. It has been shown that once children turn 4 years old, the influence of nutritional interventions on height is substantially reduced (Martorell and Habicht, 1986).

Based on this evidence, even though we could expect Oportunidades to potentially improve the nutritional status of all children, the improved nutritional status should only be manifested on height if children were young enough at the time Oportunidades arrived. This logic is what constitutes the basis for the definition of treatment we follow in the study: treated children are defined to be those who are younger than 48 months at the time the program arrived, and they are exposed to the program until they reach that age. Older children constitute our control group.[25]

Note that up to this point we have not mentioned at all actual program participation. In our baseline specification, we only define exposure as a function of age, place of residence, and time of measurement. This means we follow an intent-to-treat approach, where every child in a given place will be treated as exposed if that child lives in a community that was covered by the program while the child was younger than 4 years old. Later in the section we come back to this point. For now, we want to make two points. First, in the empirical analysis we complement this baseline specification with other specifications that, while keeping the definition of exposure, restrict the analysis to population groups more likely to participate and therefore benefit from the intervention. We use alternative variables to stratify the sample in different ways. Second, we present later a second identification strategy for which we will be able to use actual participation.

Now we explain the role of the second component in our identification strategy – the roll-out of the program. The geographic expansion of Oportunidades was far from random. As mentioned in Section 1.2, Oportunidades used census information to select which localities would receive the program first and which ones would receive it later.

---

[25] Throughout this section exposure is always defined in terms of height. Older cohorts clearly benefit from the program in other dimensions.

The stated targeting policy was to cover more disadvantaged areas first. Given this policy, we chose not to use as our control group children from places where the program had not arrived by the time of measurement. The reason is that it is very unlikely that controlling for observable locality characteristics will capture all of the locality-specific effects on height. Instead, we use the timing in the roll-out of the program to define relatively homogeneous groups of localities and perform an impact analysis within these locality groups. As will be explained below, we do use, however, children in communities where the program does not operate to control for nation-wide time trends. Finally, dividing localities in this way makes it easier to compare our results with those found in the literature. We do this in Section 1.8.

Using within-locality variation means that we need to identify impact effects from differences in the level of exposure between younger and older children. Transforming height into z-scores helps controlling for age-specific differences. However, we might still worry that the dynamics of height over the life course could contaminate the comparison between older and younger children at a given point in time. The third component of our identification strategy specifically addresses this concern. By measuring children at two points in time, 2002 and 2005, we are able to compare different cohorts of children but measured when they are at the same point over their lifecycle.

Having introduced the key elements of the identification strategy, we now explain in more detail what exactly it is that we do. To that end, we use the Lexis diagram shown in Figure 2. The horizontal axis measures time in years and the vertical axis measures age in years. The vertical line on the left is placed at the time Oportunidades first started, 1997. The other two vertical lines, at 2002 and 2005, correspond to the years MxFLS measures height.

Taking into account our definition of exposure and the time of our survey, we divide our sample of children in three groups: a younger cohort, a middle cohort and an older cohort. The diagonals in Figure 2 separate these three groups. The younger cohort consists of children between 1 and 3 years old in 2005 (born between 2001 and 2003). The middle cohort includes children 1 to 4 years old in 2002 and between 4 and 7 in 2005 (born between 1997 and 2000). The older cohort includes children between 5 and 7 years old in 2002, born between 1994 and 1996. We do not want to include children much older in the analysis, because once children reach puberty differences in height are hard to interpret. As a result, four cohort-time groups are used to identify treatment effects: the older and middle cohorts measured in 2002 and the middle and younger cohorts measured in 2005.[26]

---

[26] In the estimation only a subgroup of the middle cohort is used in order to compare groups of children as similar as possible. More specifically, since the younger cohort in 2005 will be compared to the middle cohort in 2002, only children between 1 and 3 years old in 2002 are included in the analysis. Similarly, the middle cohort in 2005 will be compared to the older cohort in 2002, so only children between 5 and 7 years old in 2005 are included in the analysis. These subgroups of the middle cohort are highlighted in Figure 2.

Next, localities are divided into different groups depending on the year they were incorporated into the program. As mentioned before, the rationale behind this lies on the fact that Oportunidades followed a specific geographic-targeting policy to incorporate new localities over time. Taking this into consideration together with the fact that height is measured in 2002 and 2005, we define four groups of rural localities and three groups of urban localities. In the rural sector the groups are the following: the first group consists of localities that were incorporated at the very beginning of the program, in 1997 or 1998; the second group consists of localities incorporated right after the first group, between the years 1999 and 2002; the third group includes the localities that received the program between 2003 and 2005; and finally the fourth group includes the localities that either received the program after 2005 or never did. Because the program only expanded to cover urban areas in 2001, only the last three groups of localities exist in the urban sector.[27,28] Note that the selection of the groups is closely related to the two years in which height measures are taken. In other words, the three groups of urban localities correspond to those incorporated to the program up to the time the first measure of height was taken (2002), those incorporated between the two years measures were taken (2002-2005) and those without Oportunidades by 2005. Rural

[27] Community group 2 is not exactly the same in the rural and urban sectors, because in the urban sector this group includes three localities incorporated to Oportunidades in 1998 (See Figure 1). In terms of interpretation and exposition, however, they should be thought of as the same thing.

[28] Results are robust to reclassifying the urban localities introduced to Oportunidades before 2001 as rural (See footnote 23).

localities have an additional fourth group that corresponds to those localities incorporated to the program when Oportunidades had just begun (1997 and 1998). Additionally, this first rural group constitutes a benchmark for comparison with current evidence, as the rural evaluation sample represents localities incorporated in 1998.[29] These groups and the time of incorporation are shown at the bottom of Figure 2.

At the time of measurement, the three cohorts of children will have experienced a different level of exposure to the program depending on the locality group they live in.[30] To identify impact effects we are going to define three levels of exposure: Full Exposure refers to children exposed to the program since birth; Partial Exposure refers to children exposed to the program for some time before the age of 4 but not since birth; and Zero Exposure refers to children that were too old at the time Oportunidades arrived (4 or older).[31] For instance, Figure 3 shows the expected level of exposure for each cohort of children if we focus on the first group of communities. At the time Oportunidades arrived, the older cohort was between 1 and 4 years old. Therefore, we define this group of children to be partially exposed when we measure them in 2002. In contrast, both the middle and younger cohorts were born either at the time the program

---

[29] Technically, the experimental evidence corresponds to communities incorporated in 1998. Longer-term non-experimental estimates include localities incorporated in 1998 and 1999.

[30] Implicit here is the assumption that the place of residence of our children at the time of the survey is the place of birth. Otherwise, the date of incorporation of the locality of residence into the program could not be used to determine level of exposure. This should not be a major concern if we think it is mostly young and single individuals who move.

[31] In this section we refer to full or partial exposure broadly defined. In the empirical section we quantify potential exposure in months for each cohort in each locality group. This will help interpret the results.

arrived or after Oportunidades was already in the community. As a result, these two groups of children are fully exposed to the program. The same information is presented in a table form in Table 1. Each cell represents the expected level of exposure to the program for a given cohort in a given locality-group at a given time. The first column in each of the 2002 and 2005 panels reproduce the information in Figure 3 for community group 1. Following the same reasoning we fill the rest of the cells for the other two community groups. By construction, all four cohort-time groups in community group 4 experienced zero exposure. Looking at the cells in this table, we can already see the within locality-group variation that we can exploit to identify program effects.

Based on the given definition of cohorts and locality groups, we estimate the following regression equation:

(1.1) $\theta_{icvt} = \alpha_{ct}^{4} + \alpha_{ct}^{1} I_1 + \alpha_{ct}^{2} I_2 + \alpha_{ct}^{3} I_3 + \gamma'_c X_{ivt} + \varepsilon_{ivt}$,

where θ_{icvt} represents the height-for-age z-score of individual i, in cohort c, in community group v, at time t. The superscript on each α coefficient represents the locality group.[32] The specification allows for four different cohort-time intercepts (α_{ct}^{4} terms): an intercept for the middle and older cohorts at time zero (year 2002), and an intercept for the middle and younger cohorts at time one (year 2005). They estimate, for a given cohort-time group, the average height-for-age z-score of children in baseline or control communities. Each of these four intercepts interacted with a dummy that

---

[32] The specification for the urban sector does not include the α_{ct}^{1} terms.

represents community-group 1, estimates, for the corresponding cohort-time group, the difference in average height-for-age of children in communities that received the program between 1997 and 1998 relative to children in baseline communities (α_{ct}^{1} terms). For instance, α_{o02}^{1} measures the difference, in 2002, between the average z-score of children in the older cohort in community-group 1 and the average z-score of children in the older cohort in baseline communities. α_{ct}^{2} and α_{ct}^{3} are interpreted in a similar manner. The set of covariates X includes: gender, age in months, presence of mother and father in household, height of mother and father, education of the mother, and state of residence. Note that the effect of each of these covariates is allowed to change across cohorts.[33]

Now we describe how we get at program effects from the estimated coefficients in equation (1.1). In the absence of selective program placement, these coefficients would already represent impact effects. But we know that is not the case. Broadly speaking, we control for program placement by using differences across α_{ct}^{v} terms for each locality group v. Because these terms are estimated relative to average z-score in control communities, these differences automatically eliminate any general time trend that makes our cohorts different from each other for reasons other than Oportunidades exposure. Finally, we control for the dynamics in height by comparing the middle cohort

[33] We tried specifications that allow for the covariates to be cohort-time specific and they essentially produce the same results (except that some coefficients have larger standard errors).

measured in 2005 with the older cohort measured in 2002 (5 to 7 years old at the time of measurement), or the younger cohort measured in 2005 with the middle cohort measured in 2002 (1 to 3 years old at the time of measurement).

This methodology leads us to the following parameters of interest: (i) community group 1, (ii) community group 2, and (iii) community group 3.

1.5.1 Community group 1: received the program between 1997 & 1998

Looking at Table 1 (or Figure 3) we see that $(\alpha^1_{m05} - \alpha^1_{o02})$ gives an estimate of full relative to partial exposure. Because both coefficients are measured relative to children in baseline communities, the difference between them controls for any locality-specific effect that can be expressed as an additive component common to both cohorts. Furthermore, we are comparing children 5 to 7 years old in both cases, which controls for any height-specific life-cycle effect. However, we are comparing different cohorts. This means that the estimate will also capture any time effect that might have existed between the time period the older cohort was in its critical growth years and the time period the middle cohort was in its critical growth years. If these communities experienced economic growth between those two periods, the middle cohort may have been exposed to a better environment than the older cohort when they both were between 1 and 4 years old. Under those circumstances, the difference between α^1_{m05} and α^1_{o02} would include both the additional exposure to Oportunidades and the improvement over time that would have happened regardless of the program.

Note, however, that only *differential* time effects between locality groups are what we need to further control for. If time trends between community groups 1 and 4 were parallel, the difference between these two cohorts in control communities would take care of any relevant time effect. We can directly test the importance of this effect if we are willing to assume that the time effect is homogeneous across cohorts. The reason is that we can use the difference between the younger cohort in 2005 and the middle cohort in 2002, as these are two groups of children that experienced the same level of exposure to the program by the time of measurement and live in the same group of communities. As a result, $(\alpha^1_{y05} - \alpha^1_{m02})$ provides us with an estimate of the time effect between the middle and younger cohorts in community group 1 relative to the same time trend in community group 4.

Assuming the relative difference across communities between the middle and younger cohorts is a good indicator of the relative difference between the older and middle cohorts, the double difference $(\alpha^1_{m05} - \alpha^1_{o02}) - (\alpha^1_{y05} - \alpha^1_{m02})$ should give us an unbiased estimate of the program effect under full exposure relative to partial exposure.[34]

[34] Note that the identification assumption regarding time effects is very specific. We need that the difference in the environment between the time period the middle cohort was 1 to 4 years old and the time period the younger cohort was 1 to 4 years old, be a good counterfactual for the difference between the time period the older cohort was 1 to 4 years old and the time period the middle cohort was 1 to 4 years old. The time effect between 2002 and 2005 is not the relevant concept because the time trend between those years affects the height of the three cohorts differently.

1.5.2 Community group 2: received the program between 1999 & 2002

Table 1 tells us that the difference $(\alpha^2_{y05} - \alpha^2_{m02})$ provides with an estimate of full relative to partial exposure, and $(\alpha^2_{m05} - \alpha^2_{o02})$ gives an estimate of partial exposure. In this case, we cannot separately estimate differential time trends between community groups 2 and 4. As opposed to the other two groups of communities, we do not have in this case two cohorts of children that experienced, at the time of measurement, the same level of exposure to the program. Consequently, we have to interpret the estimates we show in the following section as capturing both program effects as well as any possible differential time effects that could exist between these two community groups.

1.5.3 Community group 3: received the program between 2003 & 2005

Following the same reasoning explained so far, we can get an unbiased estimate of partial exposure for children in the latest community group. $(\alpha^3_{y05} - \alpha^3_{m02})$ estimates the combined effect of partial exposure to Oportunidades and differential time effects between locality groups 3 and 4. By comparing two groups of children that were not exposed to the program, $(\alpha^3_{m05} - \alpha^3_{o02})$ provides an estimate of differential time trends. Under the assumption that the time effect between the older and middle cohorts is a good counterfactual to the time effect between the middle and younger cohorts, the double difference $(\alpha^3_{y05} - \alpha^3_{m02}) - (\alpha^3_{m05} - \alpha^3_{o02})$ gives an estimate of partial exposure to the program net of placement and time effects.

1.5.4 Discussion

The combined use of the biology of child growth and the geographic expansion of the program to identify children exposed to the program and children not exposed to it constitutes a powerful identification strategy to estimate an intent-to-treat effect.

On one hand, we do not need to worry about the traditional self-selection bias that is characteristic of most impact evaluation analyses. In this analysis we work with a definition of exposure that is exogenous to the household. Therefore, household-specific factors correlated with both actual treatment status and the outcome of interest do not affect the estimates here.[35] Previous literature summarized above provides evidence of the importance to control for selective access and selective participation.

On the other hand, the definition of treatment at the locality level makes the analysis robust to the existence of spillover effects. There is no evaluation assessing the existence of these effects on health outcomes, but there is evidence on other dimensions. Bobonis and Finan (2006) find that the program significantly affected school enrollment rates among non-beneficiary children and Angelucci and DeGiorgi (2009) find that Oportunidades increased food consumption among non-beneficiary households.

---

[35] Implicit here is the assumption that the time Oportunidades arrived to the locality of residence is exogenous to the household, which holds if Oportunidades did not induce migration from places without the program to places with the program. Note additionally that the fact that children's height cannot be affected after they reach certain age rules out the possibility that parents compensate untreated children. If that was the case, behavioral responses induced by program participation would need to be considered even if treatment status was exogenous to the households.

Nevertheless, the analysis is not assumption-free. There are a few elements that deserve further discussion. And we turn to those now.

First of all, the definition of exposure at the locality level that helps us dealing with self-selection and spillover effects comes at the cost of defining a "treatment group" that may include a very large proportion of children that did not benefit from the program. In the analysis described so far, exposure is only determined by the age of the child, and no other socio-economic characteristic of the household is taken into account. However, the program targets poor households. Better-off households are not eligible and therefore their children are not expected to benefit from this intervention.[36]

To the extent that non-treated children cannot be made worse-off by the program, the estimated impact would provide a lower bound of the program effect on children's height. However, the analysis may end up being uninformative if there is not enough power to identify positive impact effects. This would be the case if the share of untreated children in the cohort exposed to the program is sufficiently large. The degree to which exposed cohorts were actually affected by the program will vary by locality. In places with a higher proportion of poor households and higher participation rates among eligible households, the proportion of non-treated children should be lower than

---

[36] Every intent-to-treat effect estimator includes both treated and non-treated individuals. Note, however, that the share of non-treated individuals is expected to be higher in this case than in more traditional impact estimates. The reason is that this analysis not only classifies as treated those children that are eligible but decided not to participate or children that participate but have limited access, but also children not designed to participate.

in other places. Given that Oportunidades was first introduced in more disadvantaged places, the identification strategy as explained so far is expected to be weaker as we evaluate communities incorporated later in time. This effect is reinforced by the lower take-up rates among eligible households after 2001. To address this concern, we add to the baseline specification additional analyses that follow the same identification strategy, but restrict the sample in different ways. We continue to use variation in exposure between older and younger children, but only use the sample of children that live in households that, according to different criteria, are more likely to participate in the program.

Also important to the identification of treatment effects are the assumptions regarding time trends and shocks. The underlying assumption regarding time effects comprises two components: how time trends change across cohort groups, and how time trends differ across locality groups.

Regarding the first component, if different cohorts were exposed to different shocks, the estimated impact will not be able to disentangle the effect of the program from the effect of these shocks. This source of bias is partially controlled for by estimating differences across cohorts in treatment communities relative to control communities. However, if there are location-specific shocks that affect differently our exposed and non-exposed cohorts of children, these effects are not controlled for.

We analyze two sources of information to explore whether this concern could be driving our results. We were informed that some localities (or households) were incorporated into Oportunidades because they suffered a negative shock, in most cases due to hurricanes. We were able to access the list of localities in which some households were enrolled in the program under a special process due to these kinds of reasons. Fortunately, only 3 MxFLS localities were in this list, and all results are robust to the exclusion of these localities. Additionally, we use information on past shocks as reported in the MxFLS household and community questionnaires. We further explain what we do in the following section.

A related, though different, concern has to do with a possible bias arising from location-specific time trends. The regression equation controls for state-time fixed effects, and all effects are estimated relative to control communities. However, if there is permanent divergence or convergence in growth rates between our treatment community groups and the control community group, time effects are not completely controlled for. Also, one might worry that the relatively small number of control communities in our sample might affect our results. To address this concern, we also report results following an alternative identification strategy that does not use children in control communities. We estimate equation (1.1) for each locality group separately, which means that now impact effects are directly identified out of differences across the four cohort-time intercepts.[37] The cost to following this alternative is that now we need somewhat stronger assumption to control for time trends within locality groups. For community groups 1 and 3, we need the difference between a given cohort and the younger (older) group to be a good counterfactual for the difference between that cohort and the older (younger) group. Furthermore, there is no way to control for time effects when we evaluate the second group of communities. One additional feature of this alternative is that now the analysis can be restricted to the sample of households who report being program participants.[38]

Lastly, there is a final point that deserves some consideration. Mexico experienced a financial crisis in 1994, known as the Tequila crises, which generated a widespread fall in income, consumption and wages. This is important for us because the group of children we labeled "older cohort" includes those children born between 1994 and 1996. Two important points need to be made. First of all, as mentioned before, nation-wide business-cycle effects are captured by children in control communities. Additionally, we allow for time-specific geographic variation by having state-time fixed effects in our estimation. Secondly, the evidence suggests that the greatest effect was on households living in metro-areas, with highly educated household heads, and workers in financial services and construction. In contrast, the smallest effect was on less-

[37] Due to sample size constraints covariates are no longer cohort specific.

[38] This is not possible under the main identification strategy, because by construction, there are no program participants in control communities.

educated, rural and agricultural workers (McKenzie, 2003). This suggests that our target population was relatively less affected by the crises. We should worry less about this issue when we estimate program effects for the rural sector, as well as when we restrict the urban sample to proxy the target population of this program. Still, we should keep this in mind as we read the results.

1.6 Descriptive analysis

Table 2 characterizes the final sample. The original sample of children 1 to 3 and 5 to 7 years old in rural communities is 3,500 and that of urban communities is 4,849. Some cases are dropped from the analysis due to lack of measurement, and only a few additional cases are lost due to measurement error or change of residence.[39] As a result, 86% of the rural sample and 79% of the urban sample are used in the analysis that follows.[40]

Summary statistics of the 2002 and 2005 z-scores show that children 1 to 8 years old in Mexico are, on average, 0.56 standard deviations below the reference median in 2002 and 0.42 below the reference median in 2005 (standard deviations are 1.26 and 1.47,

---

[39] Children that moved between 2002 and 2005 are eliminated from the 2005 sample because treatment status based on the place of residence does not correspond to actual treatment. See Section 1.5 for more details.

[40] Supplementary Section 1.10 analyzes in detail lack of measurement in children's height. It shows that missing data is clearly not random, but fortunately, interactions of main parental characteristics with the relevant groups suggest there are no differential effects across the cohorts or locality groups used in the analysis. Nonetheless, concerns regarding selectivity of measurement cannot be ruled out completely.

respectively). The percentage of stunted children, that is, children that are more than two standard deviations below the reference median, is 12% in both years.

As expected, the nutritional status of children in the rural sector is worse than that of children in the urban sector. Rural children are on average 0.75 and 0.59 below the reference median in 2002 and 2005 respectively, whereas the corresponding numbers for urban children are 0.42 and 0.29. With respect to stunting, the overall incidence of 12% is a combination of an incidence of stunting among rural children of 16% and an incidence of stunting in urban children of 10%.

Two other sets of comparisons between the rural and urban sectors are worth looking at in order to better interpret the results presented in the following section.

In the first place, participation rates in the rural and urban sector are very different. Two factors make participation rates much lower in the urban sector: the percentage of eligible households is smaller, and the take-up rate among eligible households is smaller. To document how participation rates look like, we combine the *Expansion* data with census data and calculate, for 2000 and 2005, the ratio of number of families in Oportunidades over the total number of households in the community. Table 3 shows the distribution of these ratios broken by locality groups. As expected, participation rates are much larger in the rural sector. On average, rural communities have between 30% and 70% of the households enrolled in the program, while the urban communities have at most 17%.[41]

Another difference between the rural and urban sectors has to do with the differences across the locality groups that we use in the analysis. Due to the geographic targeting of the program, we would expect localities incorporated later in time to be relatively better-off. To see how the locality groups look like, we put together information from the 1990, 2000 and 2010 censuses, and the 1995 and 2005 population counts. The information from 1990 to 2000 provide us with a picture of what these localities look like at the time Oportunidades determines which localities to introduce and at what time. Additionally, we added the 2005 and 2010 data to see whether there is any evidence of differential growth rates across locality groups.

We analyzed a broad set of socioeconomic variables, including: dwelling characteristics (floor and roof materials, access to basic services such as running water and electricity), asset ownership, literacy rates, school attendance, and access to health insurance. We show the results on a couple of these characteristics in Figure 4, but the same patterns hold for any variable we analyzed. As expected, all locality groups show signs of improvement over time. Notably, the rates of improvement do not look different across locality groups. In terms of differences in levels across locality groups,

---

[41] MxFLS reported participation rates in 2002 and 2005 are very similar, which should be expected as we have a population-representative sample in 2002.

the situation changes as we focus on the rural or the urban sector. Looking at the rural sector, both the mean and the median (not shown) of each characteristic improve as we move from community-group 1 to community-group 4. According to every indicator, the localities incorporated earlier are worse-off relative to communities incorporated later. Contrary to the rural sector, we don't see such differences across locality groups in the urban sector. Even though the levels move in the expected direction, the differences are negligible. On average, the three groups of urban communities look the same. We do not see this evidence as inconsistent with the geographic targeting of the program. While the program did target localities in the rural sector, that was not the case in the urban sector. In the later, the objective was to target, within localities, areas that have high concentration of poor households. Based on this policy, and the fact that the degree of heterogeneity within urban localities is much bigger than in rural places, it is not surprising that, on average, locality groups look the same.[42]

One variable that might deserve further explanation is access to health insurance. In 2003, the Mexican government introduced a program called Seguro Popular that provides with health insurance to households that do not have access to this service. The program expanded gradually over time, participation is voluntary, the only selection criteria is lack of health insurance, and the service is provided for free or at a very low

[42] Even though by looking at some socioeconomic indicators urban localities do not look, on average, better-off as they are incorporated later in time, the expansion was still not random. Therefore, we still think it is reasonable to keep working with the locality classification introduced earlier which uses year of entry to define locality groups.

cost. Fortunately, participation rates in our localities are not very high. In 2005 in the rural sector, they are between 5 and 20%. In the urban sector, they are between 2 and 7%. By 2010, participation rates are much higher. The fact that this program only offers an insurance product, when Oportunidades offers a wide set of benefits (including a sizable cash transfer) mitigates our concerns. However, results should be interpreted with the existence of this program in mind.

Finally, we translate our concepts of partial exposure and full exposure in terms of months. As explained in the previous section, the estimated impact effects in this analysis are identified from differences between groups of children that are fully exposed, partially exposed or not exposed to the program. In that section full exposure is defined as exposed to the intervention since birth and partial exposure is defined as some exposure during the critical growth years from some point after birth. To have a better idea of the degree of exposure we are dealing with when interpreting the estimates, we construct two measures of potential exposure. The first measure computes, for each child, the number of months there are between the time Oportunidades arrives to the locality (or birth if born after Oportunidades arrives) and the time of measurement or the time the child turns 4 years old, whichever comes first. Because one year of exposure between the ages of 1 and 2 could be very different than one year of exposure between the age of 3 and 4, in the second measure we calculate the percentage of life each child was exposed (again measured in months).[43]

Table 4 shows the distributions of these two measures for each cohort-time group in each locality group, by rural/urban sector.[44] For instance, we see that when we estimate *"Full relative to Partial exposure"* for community-group 1 in the rural sector, we are comparing children exposed to the program for 4 years to children exposed on average for a year and a half. However, the same estimate for the second group of communities actually estimates the effect of one additional year of exposure, as it compares children exposed on average for a little over 2 years versus children exposed for a little bit more than a year. Something similar happens when we talk about *"Partial exposure"* in the rural sector. In community-group 2 it represents almost 2 additional years of exposure as opposed to a little over a year in community-group 3. In the urban sector, *"Partial exposure"* represents about 15 months of exposure for both locality groups, and "*Full relative to Partial exposure*" in locality group 2 represents two additional years of exposure.

[43] These measures are only approximations since we don't have the exact month Oportunidades arrived to each place. If Mode Year data and Expansion data coincide in the year, we use the month of incorporation recorded in the Expansion data. Otherwise, we use 6.

[44] Differences in exposure between rural and urban children mainly result from differences in the distribution of the year of entry to Oportunidades within each locality group. To a less extent, it may come from individual variation in birth date and interview dates.

1.7 Results

Table 5 shows estimated coefficients on a selective group of variables of equation (1.1).[45] It is divided in three panels. The upper panel is a copy of Table 1, and presents the expected impact of the program as a function of the cohort of the child, the year height was measured and the type of locality in which he/she resides. Panel B below shows the estimated results for the rural sector, and Panel C at the bottom shows the results for the urban sector. Each cell presents the estimated coefficient corresponding to the α_{ct}^{1}, α_{ct}^{2}, and α_{ct}^{3} terms in equation (1.1). In Panel B, for example, the cell corresponding to the older cohort in community-group 1 in 2002 (-0.52) shows the estimated value of α_{o02}^{1}, when equation (1.1) is restricted to the rural sample.

As mentioned before, these coefficients cannot be directly interpreted as program effects, but their sign and significance provide some evidence on the existence, or lack of, selection in program placement. Looking at the rural sector, results are consistent with the fact that Oportunidades was first implemented in the poorest communities of Mexico, and reinforces the descriptive analysis done with census data on a broader set of socioeconomic variables. As expected, rural communities incorporated to Oportunidades before 2005 are statistically different from communities incorporated later (or never incorporated) in terms of average children's height: α_{o02}^{2} and α_{o02}^{3} are negative and significant at the 5% and 10% level, when they should be zero in the

---

[45] Standard errors are clustered at the locality level.

absence of selection in program placement. This analysis cannot provide direct evidence of placement selection for the first group of communities but the literature documents they are among the poorest communities in Mexico. As a result, we see that Oportunidades was introduced first to communities where children's nutritional status was worse off, and a simple comparison between children exposed and not exposed across communities will not provide unbiased estimates of program effects.

Contrary to the rural sector, in the urban sector there is no evidence of selection in program placement at the locality level. The estimated coefficients of the older cohort at time zero are both close to zero and statistically insignificant. Once again, this is consistent with the geographic targeting policy described previously, as well as with the descriptive evidence performed with census data.

Following the identification strategy presented in the previous section, we now present the estimated impact effects of the intervention on child nutrition. We focus first on the rural sector, and then on the urban sector.

1.7.1 Rural sector

Table 6 presents the results. It is divided in three panels. The upper panel presents the estimates for the first group of rural communities, the middle panel corresponds to community-group 2, and the bottom shows results for the last group of rural communities.

The first column in Table 6 presents the estimated results of our baseline specification, where we use all children in the rural sector. Looking at the poorest group of rural communities, the evidence suggests that Oportunidades did have a positive effect on children's height. The first row shows the difference in height between the middle cohort measured in 2005 and the older cohort measured in 2002, which would reflect the impact of being fully exposed to the program relative to being partially exposed to it, in the absence of differential time effects between this group of communities and control communities. If we look at the sample statistics discussed previously in Table 4, we see that the middle cohort was, by 2005, not only exposed to the intervention since birth, but it was also, on average, exposed for over two years more than the older cohort. Given the size of the difference in the degree of exposure, and the fact that these are very poor communities, it is not surprising that the estimated effect is positive and highly significant. Next, we look at the estimated time effect shown in row 2. The negative and statistically significant estimate says that the rate of improvement over time that took place in the first (and poorer) group of communities is slower than that in control communities. Taking this into account, we get a higher estimate of "Full relative to Partial" exposure. According to this analysis, children fully exposed to the program are on average 0.88 standard deviations taller than children only partially exposed to it. This magnitude represents, for example, 3.7 cm for a four-year-old boy.

As explained in the previous section, an estimate of program exposure that is not confounded with differential time effects cannot be identified for the second group of rural communities. Keeping this in mind, we don't see evidence of a positive impact effect on these children. Likewise, we don't see any positive effect on the nutritional status of children living in the group of rural communities incorporated at last. The estimated effect of partial exposure, with or without accounting for differential time trends, is negative and statistically insignificant.

As mentioned before, our baseline specification includes all children in the analysis, even though only a subsample of them is expected to benefit from the intervention. Given that participation rates are quite large in the rural sector, we don't expect this to be a big concern. Still, we might worry that results are affected by the fact that we are including in our analysis relatively-better off children. This is especially valid considering that participation rates go down as we move to localities incorporated later in time, and it is in those communities where we do not see evidence of positive impacts. Lastly, it might be that average estimated effects are hiding a great degree of heterogeneity, and we might want to see whether results are different when we focus on more disadvantaged groups.

For all these reasons we restrict the sample in different ways and show the results in columns (2) to (4). Column (2) presents the estimates we get when we exclude from the analysis better-off children measured by their mother's education - those whose mothers have completed high-school education or more (12 or more years of education). In column (3) we further restrict the sample to those children whose mothers have at most primary education. Finally, in column (4) we implement the Oportunidades eligibility criteria and only keep the sample of children who live in program-eligible households. That is, we only keep children living in households whose poverty score is above the official program cut-off.

Looking at the poorest group of communities the story remains unchanged, except that now impact estimates are more imprecisely estimated when using mother's education to select the sample. Column (2) shows the exact same point estimate we get in our baseline specification, but the standard error is larger and the effect is not significant. If we look at column (3), the point estimate is smaller and insignificant. However, if for efficiency reasons we restricted the -0.05 time effect to be zero, the results would suggest a positive and significant impact on these children. When we focus on program-eligible children, the effects are bigger and highly significant.

Next we analyze the second group of communities. Since this group is relatively better-off than the first one, we might expect to see some differences once we focus on more disadvantaged children. Results in columns (2) to (4) confirm this intuition. The table shows that the estimated effect of partial exposure is now larger and statistically significant. This suggests that, unless differential time effects for this community group were positive and sufficiently big, Oportunidades seems to have had a positive effect on these children's height.

In contrast, the bottom panel suggests that conclusions do not change when we analyze communities incorporated after 2002. Regardless of the sample stratification used, there is no evidence of a positive impact on these children.

Even though impact effects are not statistically significant, point estimates for the last group of communities are negative and quite large. Therefore, we looked for evidence on negative shocks that might be behind these estimates. As mentioned in the previous section, we got from Oportunidades records a list of localities in which some households were enrolled in the program through a special process after having been affected by a natural disaster. This would bias our estimates downwards, as there would be perfect correlation between program entry (positive shock) and the negative shock. It turns out that only 3 MxFLS localities are in this list, and all results are insensitive to including them or not.

Additionally, we looked for evidence on negative shocks within our survey. Both the community and the household questionnaires have a section that asks about specific shocks experienced by the community (household) during the 3 years (5 years) previous to the interview. At the locality level the enumerated shocks include: flood, earthquake, landslide, fire, hurricane, drought, plague, frost and hailstorm. At the household level, the questionnaire asks about the death or illness of household members, unemployment, and crop, production animals or property losses. This information is not complete, as we still lack information on self-reported shocks at the time the older cohort was going through its critical years (0-4 years old). Still, information reported in 2002 is relevant for the middle cohort, and information reported in 2005 is relevant for the younger cohort. This is important considering the only negative estimated effect corresponds to the last group of rural communities, where the main impact is identified from differences between the younger and middle cohorts.

A simple descriptive analysis that looks at the differences across years within locality groups does not provide clear evidence that any cohort of children is particularly worse-off or better-off. However, we did notice using the household questionnaire that for some shocks, a few localities were placed far away in the right tail of the distribution. For instance, in a few localities the percentage of households that reported having lost their house, business or total crop production was as high as 30 to 40%. To confirm that our results are not driven by any of these particular places we run the regressions without these localities and all results remain unchanged. We also run the models using the entire sample but adding a dummy variable that takes the value one if the community questionnaire indicates that the locality suffered any shock in the previous three years, and that dummy interacted with each locality group. Once again, none of the results change either in the rural or urban sectors.

To close the analysis for the rural sector, we now examine estimates provided by our second identification strategy. Columns (5) through (9) show impact effects when equation (1.1) is run for each locality group separately. Once again, we first use all the children in our sample, and then we restrict the sample in different ways. We implement all the selection criteria mentioned before and we add one additional sub-sample. In column (9) we show the estimates we get if we only use children in households who self-report being in the program at the time of the survey. We can only do this for the first two groups of communities, because by construction, there are no children in participating households in community-group 3 in 2002. For the same reason we do not use this sample stratification when using the baseline specification – there are no children in participant households in control communities.

The message is remarkably similar across columns. For the poorest group of communities, point estimates are around 0.17 and 0.28. Even though net estimates of exposure are not statistically significant, except for column 6, they would if we restricted the small point estimates of time effects to be zero. For the second group of communities, the estimated impact of partial exposure remains positive, and for the most part significant. In addition, now the estimated impact of full relative to partial exposure is larger and significant. Finally, we still do not see evidence of a positive impact on communities incorporated after 2002.

1.7.2 Urban sector

Now we turn to the urban sector. Recall that the situation in the urban sector is quite different than that in the rural sector. On one hand, heterogeneity in socio-economic status is much higher, and the percentage of program-eligible households much lower. Additionally, the evidence suggests that participation rates among eligible households are much lower. Based on this evidence, we do not expect our baseline specification, which uses all the sample of children, to be very informative. Column (1) in Table 7 confirms this intuition. Contrary to what one might expect, it looks like children exposed to the program are actually made worse-off. While the estimated effect of partial exposure for children in the second group of communities is positive and significant, the other two impact effects are negative and significant. Since we do not expect the program to affect child nutrition in a negative way, we suspect this estimates are either driven by better-off households, or program exposure is correlated with other negative shocks. We analyze these hypotheses next.

Columns (2), (3) and (4) show impact effects when we eliminate from the analysis better-off households. Just as in the rural sector, the first two columns restrict the sample based on the education level of the mother, and column (4) imposes the Oportunidades eligibility criteria to select beneficiary households. As expected, restricting the sample in any of these ways substantially reduces the urban sample size. 80% of the children have mothers with less than 12 years of education and this number goes down to 35% if we only consider those with mothers that have at most primary education. When looking at program-eligible households, only 16% of the original sample remains. None of the results show evidence of a positive impact on child height in either locality group. It is remarkable how this result holds even after focusing on the most disadvantaged groups.

Regarding looking for evidence on negative shocks, we performed the same analyses mentioned for the rural sector. Neither administrative records nor self-reported negative shocks can explain the negative coefficients shown in Table 7.

To make sure results are not driven by the relatively small sample of children in control communities, by idiosyncratic effects in our control communities, or by differential time effects across community groups, we turn next to estimates that analyze each community group separately. These results are shown in columns (5) through (9), where the first column uses the complete sample, and the rest of the columns restrict the sample in the ways previously described. When using this alternative identification strategy, there is some evidence of a positive impact of the program on child height for children in communities incorporated before 2002. Table 7 shows that the estimated effect of partial exposure is positive and significant. Recall though, that when we use this identification strategy in community-group 2 no time effect is controlled for. The estimated coefficient is capturing both, the exposure to the program as well as any improvement that might differentiate the middle and older cohorts. The situation is different when we estimate the effects using all locality groups, as it is only differential time trends that are not taken into account. Finally, we still do not see any evidence of a positive impact when we evaluate the last community group. Partial exposure, net of time trends, remains negative and insignificant.

1.8 Interpretation of the results

In summary, the analysis in the previous section suggests that the program seems to be having a positive effect on child height in children living in the poorest communities of rural Mexico. However, we don't see evidence of a positive impact on rural communities incorporated after 2002. Similarly, there isn't strong evidence of a positive impact on urban children. Estimates are always negative in communities incorporated after 2002, and they are only positive (and significant) in communities incorporated earlier for the specification that does not control in any way for time effects.

In order to put these results in context, we compare now our results to those found in previous studies.

As mentioned before, the Oportunidades rural evaluation sample covers localities incorporated in 1998. Therefore, we can compare our results for the first group of rural communities to the existing evidence in the rural sector. Experimental evidence suggests that children 12 to 36 months at baseline are about 1 cm taller after one year in the program. Analyses that evaluate one additional year of exposure after both treatment and control groups were in the program, only find positive effects on children living in more disadvantaged households (1 to 1.5 cm). To numerically compare our estimates, we need to perform a series of approximations to translate the estimated effects on z-scores into cm. We choose to convert our results for a 2 year-old boy, as it lies in between the 12 to 36 month range of previous estimates. Our baseline estimated effect of 0.88 standard deviations translates into 2.68 cm. Our second specification shows estimates in the range of 0.18 to 0.28 standard deviations. A middle point, 0.23 sd., translates into 0.7 cm. Given that these estimates represent an average additional exposure of two years, under the assumption of linear effects we get an estimated impact between 0.35 cm and 1.34 cm for an additional year. As we see, this range incorporates previous results.

Even though our results cover a wide range of values, our preferred specification for this locality group is the one that uses all children and all locality groups in the estimation. By using all locality groups we can control for nation-wide time trends, and by including all children we are maximizing the sample of children in control communities without going too far away from the target population in community-group 1, as participation rates are very high and variation in socio-economic conditions is low in this locality group. If we take this as our main result, our impact effect is somewhat bigger than previous evidence. Among other things, this could be due to the fact that our cohort of children fully exposed to the program was, not only exposed since birth, but also exposed for 4 years.[46] Therefore, this group of children experienced the maximum level of exposure one could get from this intervention, including pre-birth exposure for some of them. Additionally, it could be that the documented problems in the initial implementation of the nutritional component of the program were fixed over time, or that there is a learning process happening at the household or community level. We are measuring impact effects seven years after the program started.[47]

The only existing evidence from the urban sector analyzes the effect after two years of exposure on children incorporated in 2002. The study does not find any effect when evaluating all the sample of children (<24mo at baseline), but positive effects on height for children who were <6 months at baseline (0.41 z-score). When the authors divide the sample by socioeconomic status, they find a positive effect (0.27 z-score) on children in the lowest SES tercile. The closest estimate we have to compare this result with is our "*Full relative to Partial exposure*" estimate in community-group 2. In that case, we are approximately estimating the same length of exposure – this estimate represents an average of two additional years of exposure. Additionally, since children exposed for longer are in the program since birth, they can be closer compared with the subsample of children for which previous estimates found positive effects (<6mo). Finally, even

---

[46] This exposure is not on average. Every child in this cohort was exposed for four years, as virtually all are born after the program is in place, and they are all measured in 2005 when they are over 5 years old.

[47] Another difference could come from sample composition effects. The Oportunidades data sampled households in 1997, while our survey represents the population of these localities in 2002.

though we are comparing children fully exposed to children partially exposed (as opposed to children not exposed), our middle cohort of children was on average, only exposed for about 4 months. Table 7 shows that this estimate is never significant, even when we focused on more disadvantaged households – most of the time is negative, and when positive the point estimate is quite small. Our estimates for the last group of communities also uses children exposed since birth, although for a shorter period of time – on average, about 15 months. These estimates do not change as we focus on more disadvantaged groups.

Probably, the biggest limitation of this work is that we cannot focus too narrowly on specific subgroups due to sample size constraints. Actually, our estimates are in general poorly estimated, especially in the urban sector, as our point estimates are not very stable and standard errors are quite high. Therefore, it could be argued that there might be positive effects on beneficiary households that this analysis is not capturing. Even though that is entirely possible, it should be noted that the only evidence we have in the urban sector implements standard non-experimental methods, is not representative of the beneficiary population incorporated in 2002, it has to deal with very high attrition rates, and only finds effects on very specific subgroups of the beneficiary population. We want to emphasize that the reason this estimation strategy is arguably weaker in the urban sector is due to the very low participation rates, and this is precisely the biggest concern with current estimates, as self-selection into treatment potentially constitutes an important source of bias. All in all, we conclude that although it is possible the program has a positive impact on child height, up to this day there is no conclusive evidence that supports this.

An alternative claim that could be made is that we are only looking at one marker of nutrition - height. Then, it could still be the case that the program has positive effects on other nutritional indicators. While this is entirely possible, previous evidence points at an ever weaker effect of the program on other outcomes. Leroy et al. (2009) summarize the evidence on nutrition for 5 CCT programs in Latin America, including Oportunidades. They show how anthropometrics have been found to be the place where CCTs have the highest impact, and within anthropometrics, height has been found to be more responsive than weight. Evidence on micronutrient status summarized in that paper shows modest impact of Oportunidades on hemoglobin and anemia prevalence, and no impact on other micronutrients such as vitamin A, iron or zinc.

Assuming we believe the program does not have an effect on height on children living in rural communities incorporated after 2002 and in urban places, finding the reasons why this is the case is not straightforward. Even though these communities are relatively better-off, we are not looking at an outcome for which there is no room for improvement. Therefore, we should expect the intervention to improve the nutritional status of participating children.

As we mentioned in Section 1.2, implementation problems have been documented both in the rural and urban sectors. There is evidence that nutritional supplements were not consumed as indicated, both in terms of frequency as well as preparation. Supplements were shared with other household members, and the meetings were not effective at changing households' habits. Additionally, it was discovered that the iron in the supplement was not easily absorbed. For all these reasons, a series of changes were introduced later: it was better stressed the need to target supplements to intended children, the formula for the supplement was improved (for better iron absorption and other micronutrients were added), and the way community meetings were held changed (to stimulate more participation as opposed to being a lecture) (Neufeld et al., 2005). These changes were incorporated after September 2005, so they are outside our sample period. Therefore, poor implementation could partially explain the lack of impact. However, to explain the results these problems need to be a bigger issue in communities incorporated later. Also, given the size of the monetary transfer, one would expect some effect even if the nutritional component is not perfectly implemented. The fact that the transfer represents a smaller percentage of expenditures for urban households might also help explain the results.

1.9 Conclusion

Oportunidades is an ambitious antipoverty program that has been operating in Mexico since 1997. At first, the program only covered poor rural areas, and the implementation of the program was complemented with an evaluation designed with the purpose of assessing the impact of the program and informing policy. The success in short-term evaluations led to a massive expansion of the intervention throughout the country, and started a new trend in the design of anti-poverty programs throughout the world. By now, about one quarter of the Mexican population is enrolled in this program.

Behind the expansion of the program is the idea that the positive impacts found on communities incorporated at the beginning of the intervention would carry over to the rest of the country. However, the program followed a very specific geographic targeting rule, covering more disadvantaged places earlier. Furthermore, the program only covered rural areas at first, and the most rigorous evidence, especially on child nutrition, come from evaluations in the rural sector. This suggests there is a very important question unanswered in the literature, which is to see whether the program is successful in improving the well-being of beneficiary households once the program expanded to cover better-off communities and urban areas. This is especially important, considering we are evaluating a very expensive intervention.

The purpose of this work is to fill this gap in the literature for the case of child nutrition. In contrast to previous evidence, this is the first time population-representative data is used to analyze this program, which allows us to perform an impact analysis at the national level. We focus on child nutrition, as it has important long-term consequences. Using height as our marker of interest, we look at an outcome that is responsive to nutritional interventions early in life, and has a well-defined welfare meaning due to the link there is between child height, adult height and earnings. Our main source of information is the Mexican Family Life Survey, complemented with Oportunidades administrative records.

In order to isolate impact effects using non-experimental data we propose an identification strategy that combines insights from the biology of child growth, the timing in the roll-out of the program, and the panel dimension of our household survey. The nutritional literature has established that height is responsive to nutritional interventions early in life, but is pretty much predetermined once children reach the age of 4. Drawing from this evidence, we define exposure to the program as a function of the age of the child at the time the program arrives to the locality of residence, and exploit the variation in exposure generated by age differences. Next, we divide our localities in different groups based on their date of entry into the program. This allows us to control for program placement effects and time trends, and to link our results to the existing evidence. Finally, we use the fact that we measure children at two points in time, to be able to control for the nonlinearity of height over the life course.

The results presented here suggest that the program did have a positive impact on young children that live in the poorest communities in rural Mexico. These results match existing evidence that exploit the experimental design of the program during the first couple of years of operation. More specifically, experimental evidence corresponds to what we defined as community-group 1, and children measured in the data used in the existing literature correspond to what we defined as the older cohort. By evaluating our effects on a younger group of children, our results extend previous evidence suggesting that the program continued to have a positive impact on younger cohorts.

However, our results also reveal an important degree of heterogeneity in impact both across the rural and urban sectors as well as across community groups incorporated at different points in time within the rural or urban sectors.

With respect to the remaining communities in the rural sector, estimated effects suggest that the program seems to have had a positive impact in our second group of communities, those incorporated between 1999 and 2002. In contrast, there is no evidence of a positive impact of the intervention on children that live in the locality group that was incorporated much later, between 2003 and 2005, more than 6 years after Oportunidades started. Across all specifications, impact estimates are statistically insignificant (and negative).

Finally, there is very limited evidence of an improvement in children's nutritional status in urban areas. Estimated effects are never significant for those localities incorporated after 2002, even when we restrict the sample to the more disadvantaged households. The only case in which estimates are positive and significant is for the group of communities incorporated on or before 2002, but only if we use a specification that does not control for any time effects.

To sum up, we find that the program significantly improved the nutritional status of beneficiary children, but this positive effect on height is confined to the poorest localities in rural Mexico. Once the program expanded to cover relatively better-off communities, the positive impact disappears. Even though we only evaluate one marker of nutritional status, height, a summary of the existing literature suggests that CCTs have the highest impact on anthropometrics, and a much weaker effect has been found on micronutrient status. Therefore, it is unlikely that the program is having a big impact on other dimensions of child nutrition.

To conclude, trying to identify impact effects in the absence of experimental data is very hard. The experimental evidence that exists on Oportunidades corresponds to a non-representative group of Mexican communities and estimates only inform on short-term impact effects. However, the program has been in place for almost 12 years now, and it virtually covers all the country. To the extent we believe the identification strategy proposed in this analysis is working, we see that while the effects remain positive in poor rural communities incorporated early in time, the benefits are at best modest in the rest of the country.

1.10 Supplement: analysis of selectivity in the measurement of height

In this section we analyze the prevalence and selectivity of missing data in children's height. Children's height is our main variable of interest, as it is used to measure the impact of Oportunidades on child nutrition. Therefore, it is important to evaluate whether children measured and not measured are significantly different in ways that can bias the estimated effects.

Table 2 shows that around 17% of children 1 to 3 and 5 to 7 years old were not measured in each wave. However, missing rates are not constant across ages. Figure 5 shows the percentage of children with missing height by age and by year. The missing rate for children younger than 1 is very high (over 40% in both years), it decreases with age until children reach about 5 years old, it remains fairly constant around 11% for children 6 to 12 years old, and it starts going up again for children 12 years and older. The pattern is very similar across both years, although missing rates are somewhat larger in 2005 for children 2 to 5 years old. The figure shows that the average rate of 17% in the sample of children used in the analysis is a combination of a 22% on the younger group (1-3 years old) and a 13% on the older group.

Because we are comparing children in a certain age group in one wave relative to children in the same age group in the second wave, it is somewhat encouraging that missing rates do not vary significantly by year. However, it is the difference in characteristics between children measured and not measured what can affect the results. Ideally, children with missing height are simply a random sample of all children, in which case missing data would not be a problem. However, that is very unlikely to be the case.

Fortunately, the fact that missing height is not random is not by itself enough to bias the results.[48] Given the identification strategy followed here, what we need is measurement not to be selective between cohorts and across localities, so that the differences in average height that we attribute to time or program exposure is actually due to selective measurement.

We present next two approaches that will help characterize the nature of missing height for the sample used in this analysis. The first strategy uses the fact that MxFLS is a panel, so that we have children measured in both waves. The second approach is a standard regression analysis.

1.10.1 Descriptive analysis

The analysis presented here compares the distribution of height of children found in one wave but not in the other to have an idea of whether children with missing height are randomly selected from the height distribution. The objective is to provide suggested evidence but it is by no means conclusive.

1.10.1.1 Missing height in 2002

To see whether children measured in 2002 are significantly different from children not measured in 2002 in terms of height we can use the information collected in

---

[48] It can affect the interpretation of the results though. Let's say that we systematically miss children in the higher end of the height distribution across all cohorts and localities. In that case the diff-in-diff estimates will identify the impact on the average child of the subgroup of children measured, but it will not necessarily apply to taller children.

2005. Assuming that missing height in 2005 is random (in the sense that children from every point of the 2005 height distribution are equally likely to be missing), we can compare the distribution of height in 2005 of those children measured in 2002 with the distribution of height in 2005 of those children not measured in 2002. Table 8 shows the results. We test whether different quartiles of the height distribution are different between children measured and not measured in 2002, and we also show the results of the Kolmogorov-Smirnov test.[49] As can be seen, we almost never reject that the distributions are the same, regardless of how we break the sample: all children in 2002, older or middle cohort in 2002, or children in each group of localities.

1.10.1.2 Missing height in 2005

Using a similar reasoning we can see whether those children measured and not measured in 2005 are significantly different from each other in terms of 2002 height. Again, the comparison using the 2002 information is informative if we assume that lack of measurement in 2002 is random. In this case, we can break the analysis into two parts: missing height due to attrition or due to lack of measurement.

To analyze the first part we compare the distribution of height in 2002 between the group of children that was found and the group of children that was not found in 2005. Table 9 Panel A suggests that the distributions of height in 2002 are not different

[49] This test is very sensitive to differences in the tails of the distribution, but we show it to complement the tests of differences in selected quartiles.

between children in these two groups. When we break the analysis for each group of communities (results not shown), we find, for the rural sector, that the median in community group 2 and the quartile 75th in community group 3 are different. In the urban sector, we find that the quartile 90th is different in community group 3.

To evaluate selection in lack of measurement, we now restrict the analysis to those households that were tracked in 2005. The results are shown in Panel B. In this case, we do see significant differences when we analyze the older and middle cohorts together. However, when we break the analysis by cohort we see that all the differences are driven by the older cohort. If we analyze each locality group separately (results not shown), there are also some significant differences in three groups of localities, but again, all the differences are driven by the older cohort. There is no difference in any locality group if we restrict the analysis to the middle cohort. Therefore, even though it seems that the distribution of height in 2002 of children measured in 2005 is different than the distribution of height in 2002 of children not measured in 2005 (conditional on being tracked), this only holds for the older cohort. Fortunately, we do not use the older cohort measured in 2005.

To summarize the results, using the height distribution of children we do not find evidence of selective measurement in either wave. However, these tests are only suggestive, because to test selective measurement in one wave we are assuming random measurement in the other.

1.10.2 Regression analysis

Now we present more conventional regression analyses. Following the structure we've been using in this chapter, we are going to study the rural and urban sectors separately.

1.10.2.1 Rural sector

Table 10 presents the baseline specification of a logit regression where 1 represents having missing height. The omitted category is the older cohort in group 1 communities. The first block of explanatory variables include the full cohort-community group interactions, as well as child's gender and age.[50] The second block includes parent characteristics: mother's education, height and age, father's height, presence of the father in the household, and dummies for missing parent's height. The third block includes household characteristics. The variable 'score' is the eligibility score computed following Oportunidades eligibility criteria. Other variables are: access to social security benefits, assets ownership (vehicle, refrigerator, and washing machine), dwelling characteristics (floor material, sanitary service, water inside the dwelling), household size, number of children, characteristics of the household head (gender, and education), a measure of crowding (number of household members per room), per-capita wealth, and log of per-capita expenditure. Finally the last block includes state dummies.

[50] Age enters in a non-parametric way using a linear spline.

Statistically significant coefficients are highlighted in the table and they clearly suggest that lack of measurement is not random. As expected, the most influential variable is the measurement of the mother: if the mother is not measured the probability that the child is not measured increases a lot. Other important mother characteristics are height and age: children from taller and older mothers are more likely to be missing. The presence of the father in the household also increases the probability of having missing height. At the household level, the Oportunidades eligibility score is significant, suggesting that children in poorer households are more likely to be measured. Other variables such as access to social security benefits, number of people per room, and a couple of states are also significant, although a test of joint significance of all the variables at the household level cannot reject the null that the effect is zero.

Finally, children in two cohort-community groups are less likely to be missing. Although only marginally significant, the fact that these two groups are different is particularly important since we are using differences across cohorts and communities to identify impact effects. At the bottom of the table we present tests of joint significance for these subgroups of coefficients. The results suggest that the cohort interactions with the reference category (which represent cohorts in community group 1) as well as the cohort interactions in community group 2 and 3 are not jointly significant. However, children in group 4 do differ in their probability of being missing.

Because we are interested in differences across groups, we present next interactions of main explanatory variables with the groups of interest. In all the specifications we keep the explanatory variables of the baseline specification.

We only present the results on the variables that are going to be analyzed. However, no matter what interactions we include, the magnitude of the coefficients (and their significance level) of all the covariates other than the cohort-locality interactions remain fairly stable across specifications. We perform a test of joint significance of all the household-level variables and we can never reject the null that the joint effect is zero at standard confidence levels. With respect to the cohort-locality interactions shown in the first block, the conclusions are almost always the same, but magnitudes and individual significance do vary somewhat across specifications.[51] In a couple of cases none of the four groups of coefficients is jointly significant, but in most of the cases communities in group 4 are significantly different.

Results are presented in Table 11. The first three columns show the estimated coefficients when we interact mother's height, mother's age and household score with the four cohort-time groups, first one at a time and all together in model 4. Mother's height seems to be important for younger children (middle cohort in 2002 and younger cohort in 2005). However, the effect does not differ across the cohorts that we compare

[51] In particular, magnitudes and standard errors increase a lot in some specifications that add locality interactions (models 4 to 8 in Table 11).

in the analysis: the test of equal effects on the older cohort in 2002 and the middle cohort in 2005 is not rejected, and neither is the test that compares the middle cohort in 2002 with the younger cohort in 2005. Mother’s age seems to be important for children measured in 2002, and the eligibility score is significant across all groups. However, the test across relevant cohorts is not rejected in these cases either.

The second panel of the table (models 5 to 8) shows interactions of the same explanatory variables by locality group. In this case we are testing whether each variable has a differential effect on the probability of having missing height depending on the locality of residence. Looking at mother´s height, we see that none of the interactions is individually significant, and they are also jointly insignificant. On the contrary, it seems that mother´s age plays a significantly different role on children that live in group 4 communities. The test of joint significance of the three interactions is rejected, although p-values are not very high. Finally, the score interacted with group 4 communities is marginally significant in one of the specifications, but the effects are never jointly significant.

Finally, we analyze the full cohort-locality interaction for each of these variables. In two cases mother characteristics are significant for children in the young cohort in community group 2, and the eligibility score is individually significant for the middle cohort in community group 3. As opposed to previous cases, now there are some groups for which the null hypothesis of no impact is rejected: the joint effect of mother's age in group 2 communities and the joint effect of the score on group 4 communities.

1.10.2.2 Urban sector

The baseline specification is the same as that used in the rural sector. Results are shown in Table 12. The fact that the mother is measured is again the main explanatory variable. Children from older mothers are still more likely to have missing height, and now also mother's education is significant. At the household level, the eligibility score is no longer individually significant. However, in the urban sector the joint effect of all the household-level explanatory variables is significant at the 5% level, and this effect remains in all the specifications presented next.

As in the rural sector, there are a couple of cohort-locality groups that have different probabilities of missing height. If we analyze the coefficients by locality groups, we see that cohorts of children in group 4 and group 3 communities are not different from each other. The interactions are both individually and jointly insignificant. However, children in group 1 communities are different: the older cohort in 2002 and the middle cohort in 2005 are significantly more likely to have missing height, and the joint effect of the four cohort-time groups is significant at the 1% level.

As we did for the rural sector, we now interact some of the main explanatory variables with cohorts and locality groups. In this case, we are going to analyze the differential effect of mother's height, age and education.[52]

Table 13 presents the results. Models 1 to 4 show interactions of mother characteristics with the four cohort-time groups. Although some individual interactions are statistically significant, none of the effects is significantly different between the older cohort in 2002 and the middle cohort in 2005, or the middle cohort in 2002 and the younger cohort in 2005. Models 5 to 8 show interactions with locality groups. Fortunately, only one variable in one model is statistically significant: mother's education has a differential effect on the probability of having missing height for children in group 3 communities relative to children in group 4 communities (although the joint effect of both interactions seems to be zero). Finally, we present the full cohort-locality interaction effects. Results suggest that there are significant differences across cohorts and localities in the effect that mother's height has on the probability of having missing height.

To summarize, the analysis suggests that missing height is not random for the group of children used in the analysis. Both in the rural and urban sectors, some parental and household characteristics affect the probability of measurement in a

---

[52] We also analyze the eligibility score, but that effect is never individually or jointly significantly different from zero.

significant way. More importantly, the probability of having missing height is not the same across cohorts and localities, after controlling for a group of individual and household characteristics. Fortunately, interactions of main parental characteristics with relevant groups show that there doesn't seem to be differential effects across cohorts or across locality groups. Fully interacted models, however, do show some important differences.

1.11 Supplement: measurement error in locality date of entry

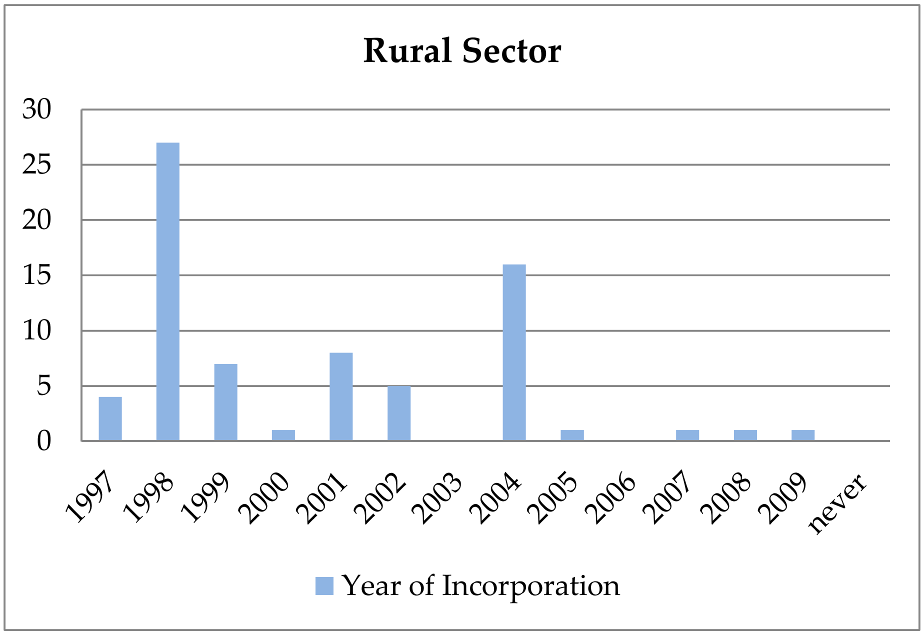
As mentioned in Section 1.4, we assign a year of entry to each MxFLS locality drawing from two sources of administrative data, complemented with self-reported participation rates in MxFLS. To make sure results are not driven by our best guess definition of the year of entry, we now show the results we get if we classify our locality groups based on each source of information, or if we only analyze the group of localities for which both data sources coincide (presumably no measurement error).

Table 14 presents the estimated results. Columns (1) and (4) show results when we use the *Mode Year* data. Columns (2) and (5) show results if we use the *Expansion* data. Columns (3) and (6) restrict the sample to localities for which the year of entry is the same according to both data sources. All these results are estimated using all children and implementing our second identification strategy – we run a separate regression for each locality group. The reason behind this is that the sample of children in control communities becomes very small if we follow the *Expansion* data.

At the bottom of the table we show sample sizes for each locality group. While there are some changes in the rural sector, we see that the biggest changes are in the urban sector. There is a big mass of children that switch from being in control communities to being in community groups 2 and 3.

The message coming from this table is that conclusions are remarkably robust across these classifications.

1.12 Primary figures



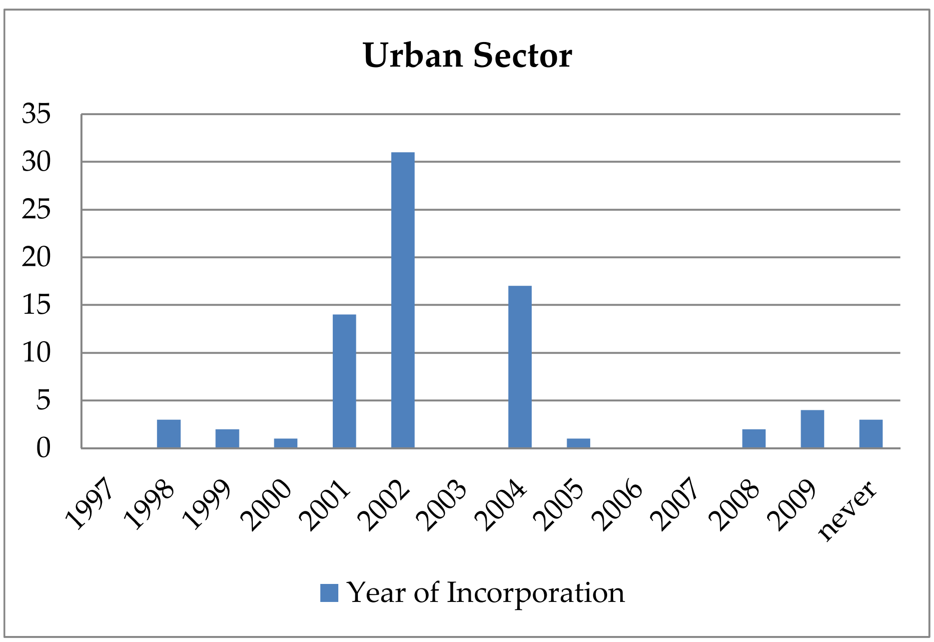


Figure 1: Expansion of Oportunidades over time at the locality level

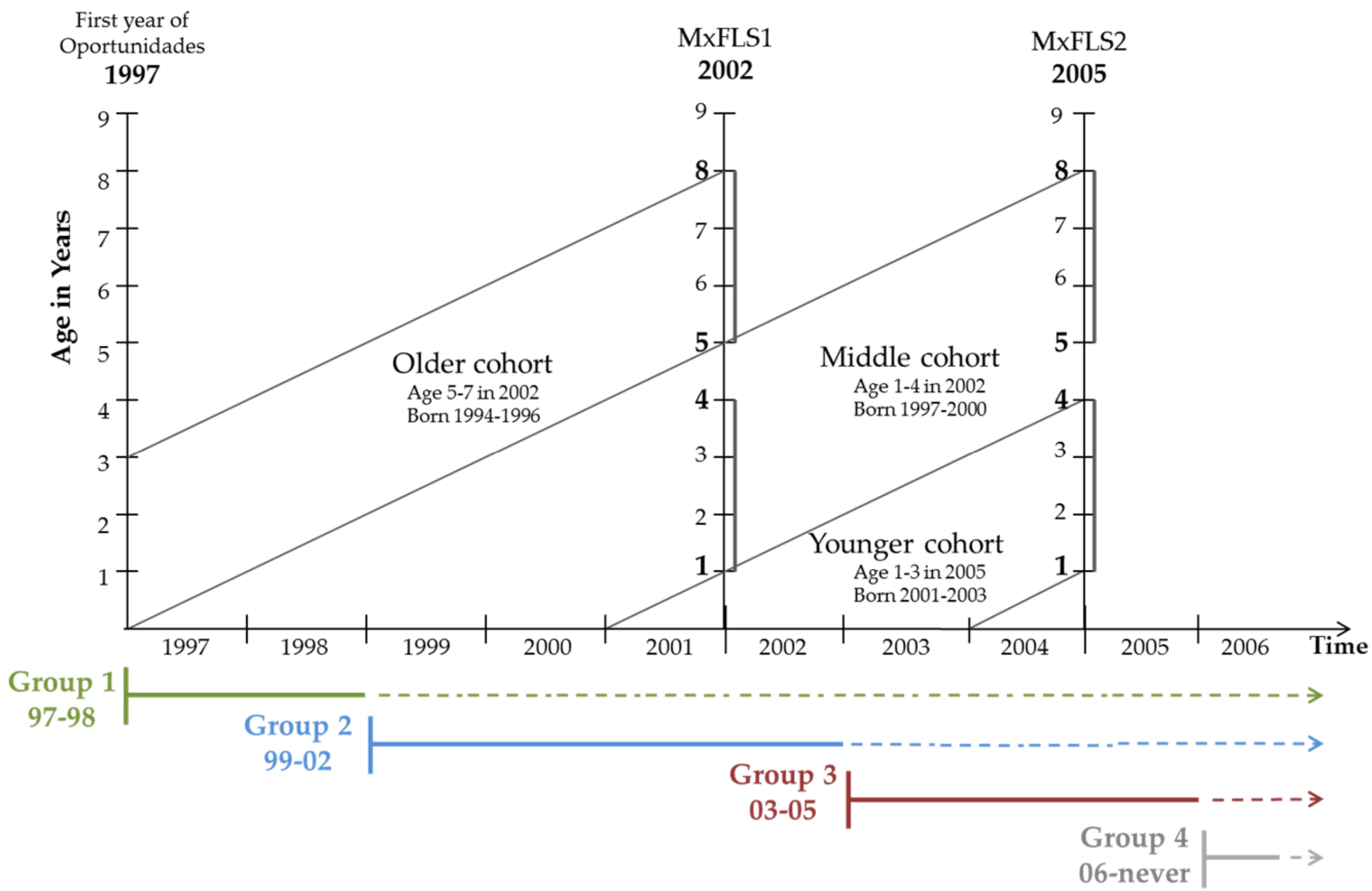


Figure 2: Graphical exposition of the identification strategy - Definition of cohorts and locality groups

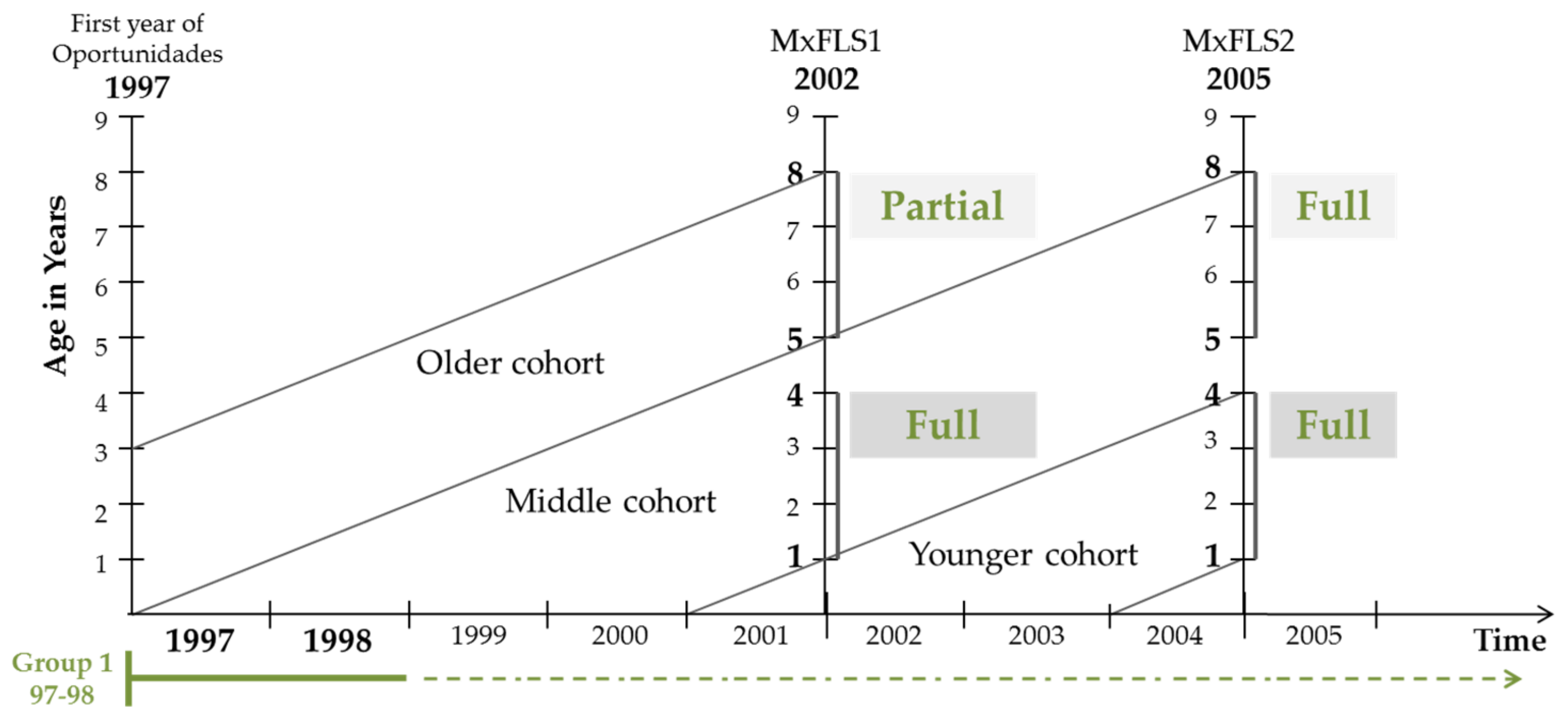


Figure 3: Expected impact of Oportunidades on children's height in Group 1 communities

1.13 Primary tables

Table 1: Expected impact of Oportunidades on child height as a function of the cohort, locality of residence and year of measurement, by rural-urban sector

Height measured in:	2002			2005		
Community group:	**G1** ≤ 98	**G2** 99 ≤ y ≤ 02	**G3** 03 ≤ y ≤ 05	**G1** ≤ 98	**G2** 99 ≤ y ≤ 02	**G3** 03 ≤ y ≤ 05
Panel A: Rural						
Old Cohort (OC)	Partial	Zero	Zero	.	.	.
Middle Cohort (MC)	Full	Partial	Zero	Full	Partial	Zero
Young Cohort (YC)	.	.	.	Full	Full	Partial
Panel B: Urban						
Old Cohort (OC)	.	Zero	Zero	.	.	.
Middle Cohort (MC)	.	Partial	Zero	.	Partial	Zero
Young Cohort (YC)	.	.	.	.	Full	Partial

Table 2: Final sample used in the analysis - Children 1 to 3 and 5 to 7 years old, by rural-urban sector

	2002		2005		TOTAL	
	#	%	#	%	#	%
TOTAL # OBS	4,542		3,807			
TOTAL RURAL	1,931		1,569		3,500	
Obs. lost because of:						
missing height	307		172			
+ missing z-score	0		3			
+ moved			19			
FINAL RURAL SAMPLE	1,624	84%	1,375	88%	**2,999**	86%
TOTAL URBAN	2,611		2,238		4,849	
Obs. lost because of:						
missing height	496		492			
+ missing z-score	0		5			
+ moved			22			
FINAL URBAN SAMPLE	2,115	81%	1,719	77%	**3,834**	79%

Notes: If children moved between 2002 and 2005 height is set to missing in 2005 (the observations are treated as if they were individuals not found in 2005)

Source: MxFLS1 & MxFLS2

Table 3: Oportunidades participation rates by locality groups

	RURAL SECTOR				URBAN SECTOR			
	p25	p50	p75	mean	p25	p50	p75	mean
COMMUNITY GROUP 1								
Incorporated between 1997 and 1998								
2000	45	57	70	58				
2005	45	58	78	64				
COMMUNITY GROUP 2								
Incorporated between 1999 and 2002								
2000	0	0	39	21	0	0	0	6
2005	40	50	73	53	4	7	17	15
COMMUNITY GROUP 3								
Incorporated between 2003 and 2005								
2000	0	0	0	0	0	0	0	0
2005	19	25	40	30	2	4	9	7

Source: Own calculations based on Oportunidades administrative records, Mexican census data and MxFLS reports.

Table 4: Potential exposure to Oportunidades, by cohort and locality group

A. Rural Sector	Months of exposure				Percentage of life			
	p25	**p50**	**p75**	**mean**	**p25**	**p50**	**p75**	**mean**
GROUP 1								
5-7 years old								
FULL - Middle Cohort 05	48	48	48	47.3	1.00	1.00	1.00	0.99
PARTIAL - Old Cohort 02	10	18.5	28	18.72	0.20	0.38	0.58	0.39
1-3 years old								
FULL - Young Cohort 05	20	28	38	28.96	1.00	1.00	1.00	1.00
FULL - Middle Cohort 2002	21	30	39.5	30.1	1.00	1.00	1.00	1.00
GROUP 2								
5-7 years old								
PARTIAL - Middle Cohort 05	13	27	41	26.14	0.27	0.56	0.85	0.54
ZERO - Old Cohort 02	0	0	3	2.95	0.00	0.00	0.06	0.06
1-3 years old								
FULL - Young Cohort 05	18	28	38	28.3	1.00	1.00	1.00	0.99
PARTIAL - Middle Cohort 2002	8	12	23	15.13	0.24	0.46	0.95	0.53
GROUP 3								
5-7 years old								
ZERO - Middle Cohort 05	0	0	0	0.6	0.00	0.00	0.00	0.01
ZERO - Old Cohort 02	0	0	0	0	0.00	0.00	0.00	0.00
1-3 years old								
PARTIAL - Young Cohort 05	13	15	18	15.23	0.38	0.56	0.92	0.60
ZERO - Middle Cohort 2002	0	0	0	0	0.00	0.00	0.00	0.00

B. Urban Sector								
	Months of exposure				Percentage of life			
	p25	**p50**	**p75**	**mean**	**p25**	**p50**	**p75**	**mean**
GROUP 1								
5-7 years old								
FULL - Middle Cohort 05								
PARTIAL - Old Cohort 02								
1-3 years old								
FULL - Young Cohort 05								
FULL - Middle Cohort 2002								
GROUP 2								
5-7 years old								
PARTIAL - Middle Cohort 05	2	14	24	15.82	0.04	0.29	0.50	0.33
ZERO - Old Cohort 02	0	0	0	1.15	0.00	0.00	0.00	0.02
1-3 years old								
FULL - Young Cohort 05	20	28	35	27.81	1.00	1.00	1.00	0.97
PARTIAL - Middle Cohort 2002	0	0	8	4.56	0.00	0.00	0.23	0.17
GROUP 3								
5-7 years old								
ZERO - Middle Cohort 05	0	0	0	0.87	0.00	0.00	0.00	0.02
ZERO - Old Cohort 02	0	0	0	0	0.00	0.00	0.00	0.00
1-3 years old								
PARTIAL - Young Cohort 05	12	13	20	15.48	0.38	0.57	0.85	0.61
ZERO - Middle Cohort 2002	0	0	0	0	0.00	0.00	0.00	0.00

Table 5: Estimated impact of Oportunidades on child height, by type of locality and rural-urban sector

Height measured in:	2002			2005		
Community group:	**G1** ≤ 98	**G2** 99 ≤ y ≤ 02	**G3** 03 ≤ y ≤ 05	**G1** ≤ 98	**G2** 99 ≤ y ≤ 02	**G3** 03 ≤ y ≤ 05
Panel A: Expected impact						
Old Cohort (OC)	Partial	Zero	Zero	.	.	.
Middle Cohort (MC)	Full	Partial	Zero	Full	Partial	Zero
Young Cohort (YC)	.	.	.	Full	Full	Partial
Panel B: Rural Communities						
Old Cohort (OC)	-0.52	-0.46	-0.28	.	.	.
	[0.092]**	[0.134]**	[0.128]*			
Middle Cohort (MC)	-0.31	-0.41	-0.27	-0.19	-0.24	-0.04
	[0.129]*	[0.162]*	[0.165]	[0.135]	[0.159]	[0.175]
Young Cohort (YC)	.	.	.	-0.860	-0.470	-0.690
				[0.216]**	[0.278]	[0.232]**
Panel C: Urban Communities						
Old Cohort (OC)	.	-0.08	0.03	.	.	.
		[0.145]	[0.146]			
Middle Cohort (MC)	.	0.35	0.29	.	0.31	0.21
		[0.119]**	[0.171]		[0.136]*	[0.175]
Young Cohort (YC)	.	.	.	.	-0.23	-0.11
					[0.229]	[0.264]

Notes: Robust standard errors in squared brackets. Reference category: communities that did not have Oportunidades by 2005. Regressions control for: gender of child, age in months, presence of mother and father in the household, mother's and father's height, mother's education, state of residence.

Source: MxFLS1 & MxFLS2

Table 6: Estimated impact of Oportunidades on child height - Rural sector

Estimation strategy	**Baseline Specification All locality groups**				**Each locality group separately**				
Sample of children	**All sample (1)**	**Mother's edu<12 (2)**	**Mother's edu<=6 (3)**	**Eligible hhs (4)**	**All sample (5)**	**Mother's edu<12 (6)**	**Mother's edu<=6 (7)**	**Eligible hhs (8)**	**In Oport (9)**
GROUP 1									
FULL relative to PARTIAL exposure (MC05-OC02)	0.34	0.41	0.49	0.72	0.17	0.17	0.20	0.28	0.25
(including time effecs)	[0.18]	[0.19]	[0.26]	[0.16]	[0.08]	[0.09]	[0.09]	[0.09]	[0.10]
TIME effects (YC05-MC02)	-0.55	-0.46	-0.05	-0.75	-0.07	-0.11	-0.02	0.04	0.07
	[0.32]	[0.49]	[0.35]	[0.58]	[0.16]	[0.16]	[0.19]	[0.15]	[0.20]
FULL relative to PARTIAL exposure	**0.88**	**0.88**	**0.54**	**1.47**	**0.24**	**0.28**	**0.22**	**0.24**	**0.18**
	[0.41]	[0.59]	[0.52]	[0.70]	[0.16]	[0.16]	[0.17]	[0.16]	[0.20]
GROUP 2									
FULL relative to PARTIAL exposure (YC05-MC02)	-0.06	0.07	0.33	-0.42	0.37	0.41	0.35	0.30	0.31
(including time effecs)	[0.37]	[0.47]	[0.34]	[0.58]	[0.16]	[0.16]	[0.19]	[0.18]	[0.15]
PARTIAL exposure (M05-O02)	0.22	0.47	0.52	0.71	0.24	0.26	0.31	0.26	0.16
(including time effecs)	[0.19]	[0.22]	[0.28]	[0.19]	[0.12]	[0.14]	[0.14]	[0.13]	[0.14]
GROUP 3									
TIME effects (M05-O02)	0.24	0.44	0.53	0.49	0.19	0.26	0.34	0.05	
(including time effecs)	[0.20]	[0.24]	[0.31]	[0.25]	[0.18]	[0.18]	[0.20]	[0.19]	
PARTIAL exposure (YC05-MC02)	-0.43	-0.27	0.21	-0.63	-0.06	0.03	0.28	0.06	
(including time effecs)	[0.32]	[0.50]	[0.36]	[0.59]	[0.17]	[0.17]	[0.22]	[0.21]	
PARTIAL exposure	**-0.67**	**-0.71**	**-0.32**	**-1.11**	**-0.25**	**-0.23**	**-0.06**	**0.01**	
	[0.44]	[0.64]	[0.58]	[0.75]	[0.30]	[0.31]	[0.32]	[0.31]	
Total Observations	3,003	2,813	1,852	2,077					
Observations Group 1					1,454	1,407	1,029	1,137	955
Observations Group 2					832	759	462	581	425
Observations Group 3					580	515	264	283	

Notes: see notes Table 5

Table 7: Estimated impact of Oportunidades on child height - Urban sector

Estimation strategy	**Baseline Specification All locality groups**				**Each locality group separately**				
Sample of children	**All sample (1)**	**Mother's edu<12 (2)**	**Mother's edu<=6 (3)**	**Eligible hhs (4)**	**All sample (5)**	**Mother's edu<12 (6)**	**Mother's edu<=6 (7)**	**Eligible hhs (8)**	**In Oport (9)**
GROUP 1									
FULL relative to PARTIAL exposure (MC05-OC02)									
(including time effecs)									
TIME effects (YC05-MC02)									
FULL relative to PARTIAL exposure									
GROUP 2									
FULL relative to PARTIAL exposure (YC05-MC02)	-0.58	-0.22	0.09	-0.23	-0.08	-0.06	0.08	-0.12	0.08
(including time effecs)	[0.24]	[0.22]	[0.38]	[0.59]	[0.10]	[0.12]	[0.19]	[0.26]	[0.36]
PARTIAL exposure (M05-O02)	0.39	0.06	0.22	-0.17	0.26	0.23	0.24	0.12	0.32
(including time effecs)	[0.19]	[0.20]	[0.26]	[0.20]	[0.07]	[0.07]	[0.10]	[0.17]	[0.11]
GROUP 3									
TIME effects (M05-O02)	0.18	-0.01	0.15	0.66	0.22	0.18	0.17	0.80	
(including time effecs)	[0.22]	#DIV/0!	[0.28]	[0.28]	[0.12]	[0.13]	[0.15]	[0.17]	
PARTIAL exposure (YC05-MC02)	-0.40	-0.18	-0.20	0.24	-0.03	0.01	-0.16	0.35	
(including time effecs)	[0.27]	[0.25]	[0.40]	[0.53]	[0.17]	[0.13]	[0.25]	[0.19]	
PARTIAL exposure	**-0.58**	**-0.17**	**-0.35**	**-0.42**	**-0.26**	**-0.17**	**-0.32**	**-0.45**	
	[0.33]	[0.30]	[0.49]	[0.66]	[0.24]	[0.21]	[0.26]	[0.28]	
Total Observations	3,828	3,160	1,331	622					
Observations Group 1									
Observations Group 2					2,585	2,147	900	467	372
Observations Group 3					835	687	285	110	

Notes: see notes Table 5

1.14 Supplementary figures

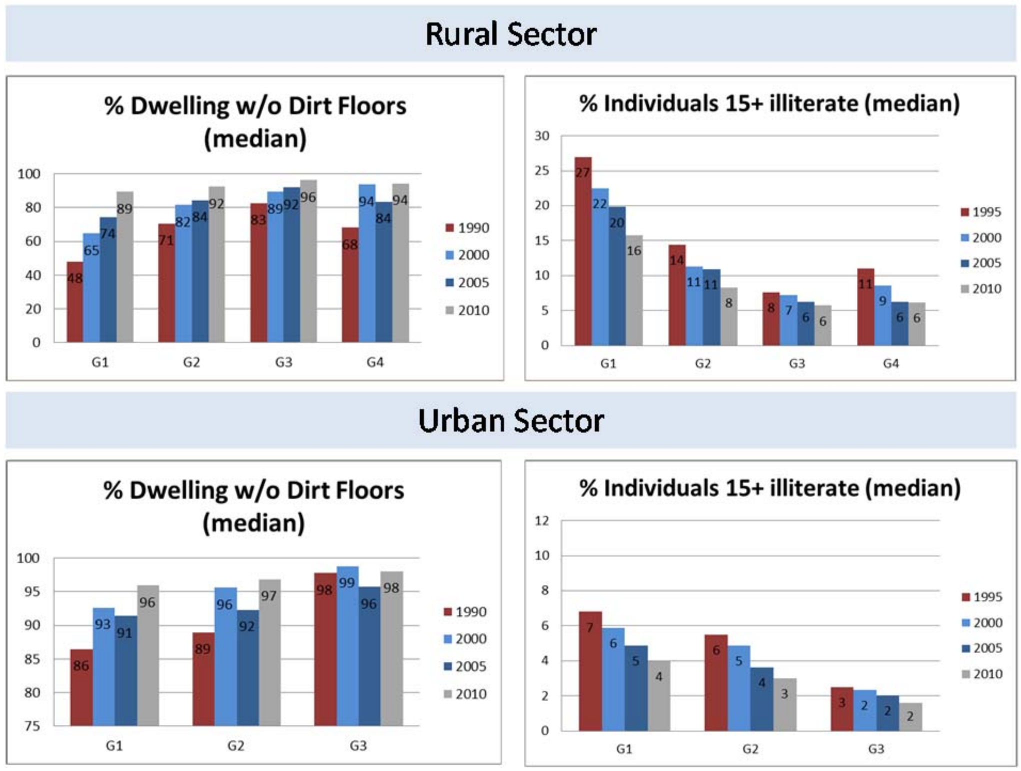


Figure 4: Description of locality groups using Census data

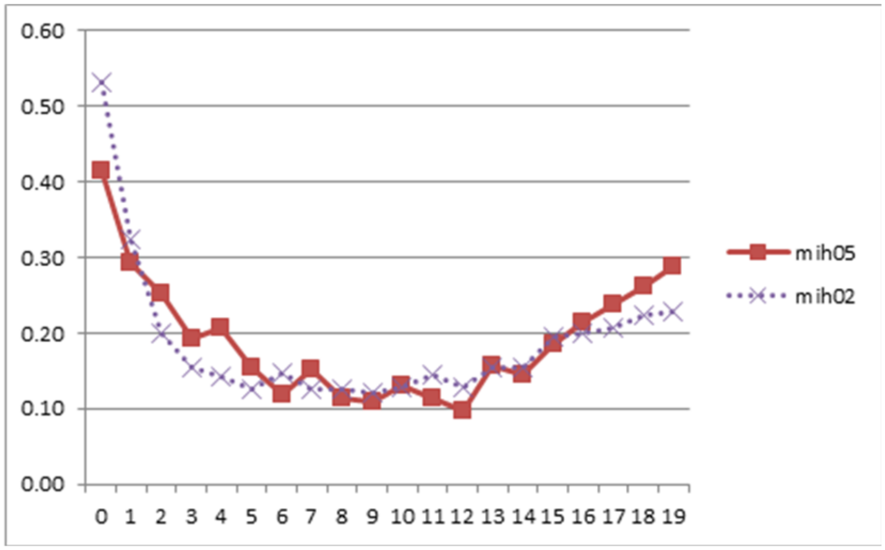


Figure 5: Proportion of children with measured height, by age and MxFLS round

1.15 Supplementary tables

Table 8: Missing height in 2002 - Test of differences between the 2005 height distribution of children measured in 2002 and children not measured in 2002

				Rural			Urban		
	OC-MC	OC	MC	G1	G2	G3	G1	G2	G4
mean	-0.01	-0.35	-0.19	-0.04	-0.23	-0.19	0.01	0.39	0.20
	[0.06]	[0.09]	[0.09]	[0.14]	[0.17]	[0.22]	[0.12]	[0.22]	[0.14]
q1	-0.12	-0.23	-0.22	-0.42	-0.24	-0.12	0.14	0.24	-0.03
	[0.11]	[0.19]	[0.14]	[0.28]	[0.32]	[0.44]	[0.20]	[0.47]	[0.27]
q25	-0.79	-0.02	-0.19	-0.12	-0.01	0.00	-0.13	0.27	0.10
	[0.08]	[0.13]	[0.12]	[0.18]	[0.19]	[0.33]	[0.13]	[0.32]	[0.19]
q5	-0.02	-0.09	0.02	0.00	-0.24	-0.26	-0.03	0.35	0.17
	[0.08]	[0.11]	[0.12]	[0.18]	[0.23]	[0.26]	[0.16]	[0.35]	[0.20]
q75	0.14	-0.06	0.20	-0.23	-0.01	-0.13	0.19	0.61	0.27
	[0.09]	[0.11]	[0.13]	[0.14]	[0.22]	[0.26]	[0.16]	[0.38]	[0.18]
q9	**0.23**	0.14	0.30	0.23	-0.33	-0.03	0.29	0.55	**0.50**
	[0.10]**	[0.15]	[0.18]	[0.23]	[0.29]	[0.37]	[0.24]	[0.44]	**[0.26]***
K-S test	**0.07**	0.09	0.07	0.07	0.09	0.09	0.10	0.15	0.10
	(0.03)	(0.113)	(0.11)	(0.76)	(0.585)	(0.93)	(0.20)	(0.428)	(0.355)
Obs w/heigth	2,949	1,670	1,279	696	378	272	739	243	549
Obs w/o height	512	195	317	113	81	42	140	35	94

Notes: Group4 in rural communities is not separately analyzed because the sample size is very small. *Significant at the 10% level, **Significant at the 5%level, ***Significant at the 1% level.

Source: MxFLS1 & MxFLS2

Table 9: Missing height in 2005 – Test of differences between the 2002 height distribution of children measured in 2005 and children not measured in 2005

	Panel A Attrition			Panel B Lack of measurement		
	OC-MC	**OC**	**MC**	**OC-MC**	**OC**	**MC**
mean	0.00	-0.01	0.02	**0.13**	**0.23**	0.05
	[0.06]	[0.09]	[0.10]	**[0.07]****	**[0.09]****	[0.10]
q1	0.01	-0.15	0.08	0.08	0.21	-0.05
	[0.12]	[0.15]	[0.19]	[0.13]	[0.16]	[0.21]
q25	-0.10	-0.06	-0.13	**0.21**	**0.30**	0.18
	[0.10]	[0.13]	[0.16]	**[0.10]****	**[0.14]****	[0.15]
q5	0.05	0.05	0.06	**0.20**	**0.32**	0.09
	[0.06]	[0.08]	[0.09]	**[0.08]*****	**[0.10]*****	[0.09]
q75	0.05	0.02	0.11	0.11	**0.20**	-0.01
	[0.09]	[0.11]	[0.12]	[0.09]	**[0.11]***	[0.12]
q9	-0.04	-0.01	-0.08	0.12	0.19	0.02
	[0.11]	[0.14]	[0.17]	[0.10]	[0.14]	[0.17]
K-S test	0.04	0.06	0.04	**0.09**	**0.15**	0.06
	(0.57)	(0.60)	(0.84)	**(0.01)**	**(0.00)**	(0.45)
Obs. w/heigth	3331	1853	1478	2949	1670	1279
Obs. w/o height	411	202	209	382	183	199

Notes: Standard errors in square brackets. *Significant at the 10% level, **Significant at the 5%level, ***Significant at the 1% level

Source: MxFLS1 & MxFLS2

Table 10: Logistic regression (=1 if missing height) - Rural sector

	Odds Ratio	Robust Std. Err.	p-value
Base: oc - G1			
mc0	0.918	0.455	0.864
mc1	0.442	0.231	0.119
yc1	0.516	0.292	0.244
G2			
oc0	1.472	0.754	0.451
mc0	0.831	0.283	0.589
mc1	1.887	0.955	0.210
yc1	1.077	0.366	0.827
G3			
oc0	0.48	0.201	0.080
mc0	0.736	0.213	0.291
mc1	1.247	0.707	0.697
yc1	1.205	0.437	0.607
G4			
oc0	0.736	0.352	0.523
mc0	0.406	0.218	0.094
mc1	0.495	0.393	0.252
yc1	1.129	0.532	0.796
Sex	0.975	0.051	0.644
Age1	0.348	0.076	0.000
Age2	0.716	0.158	0.132
Age3	0.887	0.183	0.565
Age4	0.846	0.162	0.385
moe	0.967	0.029	0.288
moh	1.021	0.011	0.057
moage	1.031	0.011	0.004
fah	1.002	0.014	0.870
fa in hh	2.692	0.659	0.000
miss moe	0.080	0.032	0.000
miss moh	28.268	7.674	0.000
miss fah	2.771	0.551	0.000
oport score	0.596	0.132	0.020
no social security	1.640	0.439	0.064
no vehicle	1.169	0.253	0.471
refrigerator/washer	1.571	0.443	0.109
soil floor	0.849	0.182	0.449

	Odds Ratio	Robust Std. Err.	p-value
no bath	1.284	0.280	0.251
no running water	0.851	0.170	0.423
no gas stove	1.609	0.479	0.110
hh size	1.052	0.047	0.257
children in hh	0.996	0.084	0.967
male head	1.273	0.285	0.280
head w/o education	1.152	0.275	0.552
head w/ primary incomplete	0.994	0.161	0.974
crowding	1.083	0.049	0.081
pc wealth	0.999	0.000	0.206
log pce expenditure	1.143	0.124	0.219
miss pcw	3.970	5.289	0.301
miss lpce	1.186	1.873	0.914
state2	1.149	0.480	0.740
state3	1.254	0.339	0.403
state4	0.967	0.304	0.915
state5	0.902	0.209	0.658
state6	1.133	0.299	0.723
state7	0.513	0.237	0.149
state8	1.143	0.259	0.556
state9	0.576	0.216	0.142
state10	1.28	0.224	0.159
state11	2.293	1.085	0.079
state12	0.522	0.138	0.014
state14	1.137	0.396	0.712
state15	1.058	0.379	0.874
state16	0.641	0.174	0.102
state17	1.453	0.694	0.434
	Chi2	**p-value**	
G1	4.89	0.18	
G2	4.64	0.32	
G3	3.01	0.55	
G4	14.3	0.006	

Notes: standard errors clustered at the locality level. Sample size: 3,436

Table 11: Logit model - Missing height for the sample of children used in the analysis – Odds ratio and Chi2 tests of key variables – Rural sector

		Cohort-time interactions					Locality interactions			
		Model 1	Model 2	Model 3	Model 4		Model 5	Model 6	Model 7	Model 8
	moh_oc0	1.00			1.00	**moh**	1.02			1.02
		[0.20]			[0.02]		[0.19]			[0.02]
	moh_mc0	1.03			1.04	**moh_G2**	1.01			1.00
		[0.18]*			[0.02]**		[0.03]			[0.03]
	moh_mc1	1.00			1.00	**moh_G3**	1.01			1.00
		[0.02]			[0.02]		[0.02]			[0.02]
	moh_yc1	1.04			1.03	**moh_G4**	1.01			1.00
		[0.02]*			[0.02]		[0.04]			[0.05]
Chi2	**(all=0)**	5.46			6.14	**G2=G3=G4=0**	0.22			0.03
		(0.24)			(0.19)		(0.97)			(0.99)
	(oc0=mc1)	0.11			0.00					
		(0.74)			(0.95)					
	(mc0=yc1)	0.08			0.02					
		(0.78)			(0.89)					
	moage_oc0		1.05		1.04	**moage**		1.05		1.05
			[0.17]***		[0.02]***			[0.01]***		[0.02]***
	moage_mc0		1.04		1.04	**moage_G2**		0.97		0.97
			[0.02]**		[0.02]**			[0.02]		[0.02]
	moage_mc1		1.03		1.03	**moage_G3**		**0.94**		**0.95**
			[0.02]		[0.02]			**[0.03]****		**[0.03]****
	moage_yc1		1.01		1.01	**moage_G4**		0.98		0.98
			[0.02]		[0.02]			[0.02]		[0.02]
Chi2	**(all=0)**		**15.06**		**14.39**	**G2=G3=G4=0**		5.54		4.88
			(0.00)		**(0.00)**			(0.14)		(0.18)
	(oc0=mc1)		0.4		0.45					
			(0.53)		(0.50)					
	(mc0=yc1)		2.11		1.17					
			(0.15)		(0.22)					
	score_oc0			0.60	0.60	**score**			0.65	0.65
				[0.13]**	[0.13]**				[0.16]*	[0.16]*
	score_mc0			0.66	0.70	**score_G2**			0.9	0.93
				[0.15]*	[0.16]				[0.13]	[0.14]
	score_mc1			0.50	0.50	**score_G3**			0.78	0.79
				[0.14]**	[0.14]**				[0.18]	[0.19]
	score_yc1			0.58	0.63	**score_G4**			**0.67**	0.7
				[0.16]**	[0.18]*				**[0.14]***	[0.17]
Chi2	**(all=0)**			**8.97**	**10.51**	**G2=G3=G4=0**			3.89	2.51
				(0.06)	**(0.03)**				(0.27)	(0.47)
	(oc0=mc1)			0.57	0.63					
				(0.45)	(0.43)					
	(mc0=yc1)			0.94	0.50					
				(0.33)	(0.48)					

Locality-Cohort interactions						
var =	Mother's height		Mother's age		Household Score	
	coeff	st. error	coeff	st. error	coeff	st. error
var	1.02	[0.03]	1.07	[0.02]	0.67	[0.18]
var_mc0	1.02	[0.04]	0.97	[0.03]	1.02	[0.14]
var_mc1	0.98	[0.06]	0.98	[0.04]	0.81	[0.30]
var_yc1	0.97	[0.04]	1.00	[0.04]	0.91	[0.25]
var_oc0_G2	0.94	[0.04]	0.97	[0.04]	0.78	[0.19]
var_mc0_G2	0.98	[0.04]	1.01	[0.03]	1.03	[0.18]
var_mc1_G2	1.04	[0.06]	1.00	[0.04]	0.91	[0.32]
var_yc1_G2	**1.09**	**[0.05]****	**0.91**	**[0.03]****	0.93	[0.28]
var_oc0_G3	1.03	[0.06]	0.95	[0.03]	1.30	[0.62]
var_mc0_G3	1.03	[0.04]	0.97	[0.04]	**0.54**	**[0.18]***
var_mc1_G3	0.94	[0.06]	0.92	[0.06]	0.72	[0.44]
var_yc1_G3	1.05	[0.06]	0.93	[0.05]	0.90	[0.37]
var_oc0_G4	0.99	[0.1]	0.97	[0.05]	0.48	[0.32]
var_mc0_G4	0.93	[0.07]	1.01	[0.04]	1.23	[0.38]
var_mc1_G4	0.93	[007]	1.00	[0.09]	0.52	[0.53]
var_yc1_G4	1.15	[0.1]	0.95	[0.05]	0.50	[0.22]
Chi2_G1	1.55	(0.67)	1.23	(0.75)	0.68	(0.88)
Chi2_G2	6.80	(0.15)	**9.71**	**(0.05)**	2.48	(0.65)
Chi2_G3	1.92	(0.75)	6.29	(0.18)	3.53	(0.47)
Chi2_G4	7.57	(0.11)	4.52	(0.34)	**24.46**	**(0.00)**

Notes: Regressions include all the explanatory variables shown in previous table. Robust standard errors in square brackets next to estimated coefficients. P-values in brackets next to Chi2 tests. *Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

In models 1 to 4 (all=0) referst to the null hypothesis that the joint effect of the four cohort-time interactions is zero, (oc0=mc1) refers to the null hypothesis that the interactions with the old cohort in 2002 and the middle cohort in 2005 are the same, (mc0=yc1) referts to the null hypothesis that the interactions with the middle cohort in 2002 and the young cohort in 2005 are the same.

In models 5 to 8 (G2=G3=G4=0) refers to the null hypothesis that the joint effect of the three locality interactions is zero

Source: MxFLS1 & MxFLS2

Table 12: Logistic regression (=1 if missing height) - Urban sector

	Odds Ratio	Robust Std. Err.	p-value
Base: oc - G3			
mc0	0.853	0.372	0.716
mc1	1.148	0.258	0.538
yc1	0.955	0.38	0.908
G1			
oc0	1.167	0.454	0.691
mc0	2.369	0.645	0.002
mc1	1.02	0.325	0.95
yc1	1.816	0.441	0.014
G2			
oc0	0.898	0.292	0.742
mc0	0.588	0.269	0.246
mc1	0.624	0.208	0.159
yc1	1.707	0.63	0.147
Sex	0.928	0.047	0.149
Age1	0.509	0.105	0.001
Age2	0.728	0.104	0.027
Age3	0.977	0.168	0.894
Age4	1.237	0.159	0.099
moe	0.968	0.018	0.091
moh	1.014	0.010	0.170
moage	1.032	0.007	0.000
fah	0.988	0.010	0.294
fa in hh	2.423	0.487	0.000
miss moe	2.250	0.570	0.000
miss moh	35.09	7.081	0.000
miss moage	0.000	.	.
miss fah	3.258	0.432	0.000
oport score	1.122	0.213	0.544
no social security	1.138	0.198	0.476
no vehicle	0.808	0.090	0.059
refrigerator/washer	0.853	0.192	0.483
soil floor	1.273	0.276	0.266
no bath	1.045	0.375	0.902
no running water	0.979	0.147	0.888
no gas stove	0.695	0.181	0.165
hh size	0.948	0.031	0.110

	Odds Ratio	Robust Std. Err.	p-value
children in hh	0.986	0.069	0.846
male head	1.146	0.223	0.483
head w/o education	0.815	0.183	0.363
head w/ primary incomplete	1.014	0.173	0.933
crowding	0.978	0.052	0.687
pc wealth	1.000	0.000	0.046
log pce expenditure	1.094	0.071	0.167
miss pcw	1.750	0.695	0.158
miss lpce	1.095	0.394	0.801
state2	1.447	0.479	0.264
state3	2.224	0.822	0.030
state4	1.162	0.443	0.694
state5	1.456	0.537	0.308
state6	1.245	0.421	0.516
state7	1.524	0.549	0.242
state8	1.216	0.441	0.589
state9	0.958	0.372	0.914
state10	2.581	1.085	0.024
state11	1.495	1.164	0.605
state12	0.958	0.409	0.921
state14	0.719	0.273	0.389
state15	0.838	0.315	0.640
state16	0.831	0.401	0.702
state17	1.412	0.642	0.448
	Chi2	**p-value**	
G1	17.9	0.001	
G2	4.62	0.329	
G3	0.74	0.864	

Notes: standard errors clustered at the locality level. Sample size: 4,566

Table 13: Logit model - Missing height for the sample of children used in the analysis – Odds ratio and Chi2 tests of key variables – Urban sector

		Cohort-time interactions					Locality interactions			
		Model 1	Model 2	Model 3	Model 4		Model 5	Model 6	Model 7	Model 8
	moh_oc0	1.03			1.03	**moh**	1.02			1.02
		[0.02]			[0.02]		[0.02]			[0.02]
	moh_mc0	1.00			1.01	**moh_b**	0.98			0.99
		[0.02]			[0.02]		[0.02]			[0.02]
	moh_mc1	0.98			0.98	**moh_g**	1.00			1.01
		[0.02]			[0.02]		[0.02]			[0.03]
	moh_yc1	1.03			1.03					
		[0.02]**			[0.01]**					
Chi2	**(all=0)**	**8.03**			**8.56**	**(G1=G2=0)**	0.73			0.62
		(0.09)			**(0.07)**		(0.69)			(0.74)
	(oc0=mc1)	2.36			2.81					
		(0.12)			(0.10)					
	(mc0=yc1)	2.17			1.80					
		(0.14)			(0.18)					
	moage_oc0		1.03		1.03	**moage**		1.03		1.03
			[0.02]		[0.02]*			[0.01]**		[0.01]**
	moage_mc0		1.05		1.05	**moage_G1**		1.01		1.00
			[0.01]		[0.01]***			[0.02]		[0.02]
	moage_mc1		1.02		1.01	**moage_G2**		1.02		1.02
			[0.02]		[0.02]			[0.02]		[0.02]
	moage_yc1		1.03		1.03					
			[0.01]**		[0.01]**					
Chi2	**(all=0)**		**19.63**		**18.9**	**(G1=G2=0)**		0.66		0.51
			(0.00)		**(0.00)**			(0.72)		(0.78)
	(oc0=mc1)		0.34		0.62					
			(0.56)		(0.43)					
	(mc0=yc1)		0.78		0.44					
			(0.38)		(0.51)					
	moe_oc0			0.98	0.97	**moe**			1.00	1.01
				[0.03]	[0.03]				[0.03]	[0.03]
	moe_mc0			0.96	0.96	**moe_G1**			0.94	0.95
				[0.03]	[0.03]				[0.03]	[0.04]
	moe_mc1			0.97	0.98	**moe_G2**			**0.94**	0.93
				[0.03]	[0.03]				**[0.04]***	[0.04]
	moe_yc1			0.97	0.96					
				[0.03]	[0.02]					
Chi2	**(all=0)**			3.22	**2.76**	**(G1=G2=0)**			3.46	3.27
				(0.52)	**(0.06)**				(0.18)	(0.19)
	(oc0=mc1)			0.01	0.12					
				(0.94)	(0.73)					
	(mc0=yc1)			0.20	0.00					
				(0.65)	(0.94)					

	Locality-Cohort interactions					
var =	Mother's height		Mother's age		Mother's education	
	coeff	st. error	coeff	st. error	coeff	st. error
var	1.08	[0.05]*	1.05	[0.04]	1.05	[0.05]
var_mc0	0.96	[0.05]	0.96	[0.05]	0.97	[0.05]
var_mc1	**0.98**	**[0.04]***	0.97	[0.04]	0.96	[0.08]
var_yc1	0.96	[0.05]	0.97	[0.04]	0.93	[0.05]
var_oc0_G1	0.96	[0.05]	0.98	[0.04]	**0.89**	**[0.05]***
var_mc0_G1	**0.95**	**[0.02]**	1.06	[0.04]	0.92	[0.05]
var_mc1_G1	**1.09**	**[0.04]**	0.98	[0.04]	0.96	[0.07]
var_yc1_G1	0.99	[0.04]	0.99	[0.03]	1.01	[0.06]
var_oc0_G2	**0.89**	**[0.05]**	0.94	[0.07]	0.91	[0.09]
var_mc0_G2	1.03	[0.05]	1.01	[0.06]	0.93	[0.09]
var_mc1_G2	1.04	[0.04]	1.03	[0.05]	0.95	[0.07]
var_yc1_G2	1.02	[0.04]	1.06	[0.04]	0.97	[0.07]
Chi2_G1	**11.13**	**(0.03)**	3.12	(0.54)	6.00	(0.20)
Chi2_G2	7.25	(0.12)	3.99	(0.41)	2.77	(0.60)
Chi2_G3	**13.49**	**(0.00)**	0.75	(0.86)	1.70	(0.64)

Notes: Regressions include all the explanatory variables shown in previous table. Robust standard errors in square brackets next to estimated coefficients. P-values in brackets next to Chi2 tests. *Significant at the 10% level, **Significant at the 5% level, ***Significant at the 1% level

In models 1 to 4 (all=0) referst to the null hypothesis that the joint effect of the four cohort-time interactions is zero, (oc0=mc1) refers to the null hypothesis that the interactions with the old cohort in 2002 and the middle cohort in 2005 are the same, (mc0=yc1) referts to the null hypothesis that the interactions with the middle cohort in 2002 and the young cohort in 2005 are the same.

In models 5 to 8 (G1=G2=0) refers to the null hypothesis that the joint effect of the tow locality interactions is zero

Source: MxFLS1 & MxFLS2

Table 14: Estimated impact of Oportunidades using different locality-group classifications

	RURAL SECTOR			URBAN SECTOR		
Locality classification	**Modeyr**	**Exp**	**Exp= Modeyr**	**Modeyr**	**Exp**	**Exp= Modeyr**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
GROUP 1						
FULL rel. to PARTIAL exp. (MC05-OC02)	0.16	0.17	0.16			
(including time effecs)	[0.09]	[0.09]	[0.10]			
TIME effects (YC05-MC02)	-0.08	-0.06	-0.07			
	[0.17]	[0.18]	[0.20]			
FULL relative to PARTIAL exposure	**0.24**	**0.23**	**0.23**			
	[0.17]	[0.18]	[0.19]			
GROUP 2						
FULL rel. to PARTIAL exp. (YC05-MC02)	0.35	0.27	0.39	-0.17	-0.07	-0.17
(including time effecs)	[0.15]	[0.14]	[0.16]	[0.12]	[0.10]	[0.12]
PARTIAL exposure (M05-O02)	0.29	0.22	0.29	0.29	0.24	0.29
(including time effecs)	[0.12]	[0.10]	[0.13]	[0.09]	[0.06]	[0.09]
GROUP 3						
TIME effects (M05-O02)	0.14	0.21	0.19	0.21	0.13	0.26
(including time effecs)	[0.17]	[0.18]	[0.18]	[0.11]	[0.12]	[0.17]
PARTIAL exposure (YC05-MC02)	0.02	-0.09	-0.04	-0.18	0.02	-0.03
(including time effecs)	[0.15]	[0.17]	[0.17]	[0.20]	[0.14]	[0.30]
PARTIAL exposure	**-0.13**	**-0.30**	**-0.23**	**-0.39**	**-0.11**	**-0.29**
	[0.29]	[0.30]	[0.30]	[0.23]	[0.21]	[0.32]
Observations Group 1	1,376	1,257	1,179			
Observations Group 2	835	1,030	757	1,784	2,751	1,784
Observations Group 3	636	601	579	655	974	389

Notes: see notes Table 5

2. Extended Families across Mexico and the United States

2.1 Introduction

Existing literature on resource allocation has almost exclusively focused on households; however, we know that households are not isolated units and that interactions among non-co-resident family members are important. In general, the economic role of the family is thought to be particularly relevant in developing settings, where the availability and/or quality of market- or government-provided goods and services are low. For example, in the absence of well-functioning financial and insurance markets, families have been found to facilitate investments and engage in informal insurance agreements, among other services.[1] Additionally, the incidence and magnitude of inter-household transfers evident in multiple data sources provides further evidence on the existence and importance of connections across related households.

Taking into account interactions among non-co-resident family members can have important implications. The impact of public interventions or the development of new markets can be quite different in a world where family agreements are prevalent relative to the impact they would have in a world without such interactions. Evidence of this, for example, comes from the body of research that evaluates the famous conditional

[1] See for example Rosenzweig and Stark (1989), Angelucci et al. (2010), Thomas and Frankenberg (2007).

cash transfer program Oportunidades, which takes place in the same setting as this work - Mexico. For instance, Angelucci et al. (2009) show very clearly how children in non-beneficiary households increased their schooling if they had family members in other households in the village who did participate in the program but not otherwise. Similarly, Albarran and Attanasio (2005) show how private transfers adjust to the existence of public transfers, a topic that has drawn a lot of attention in the literature.

While it is fairly well accepted that interactions among non-co-resident family members are important, our understanding of inter-household family behavior is still quite limited, especially relative to our understanding of intra-household behavior. In this study I contribute to the literature by analyzing the extent to which Mexican families share resources across households, looking in particular at families with different degrees of spatial dispersion among their members. To the best of my knowledge, this is the first attempt to study family behavior with various degrees of spatial dispersion in a unified framework, *including* families with members spread across international borders.[2]

Taking spatial dispersion explicitly into account is important because there is reason to believe that family behavior can be different when family members live far away from each other. In the first place, families with members spread across different

[2] In line with the literature, I take household formation and location decisions as given. In other words, taking the family structure as it is, I stratify families in different groups based on the spatial dispersion among its members and analyze allocation decisions. While family structure can be endogenous to the family decision process, extending the analysis in that direction is beyond the scope of this paper.

locations do not necessarily remain connected. In the second place, when they do stay connected, they presumably have to overcome more difficulties to sustain cooperation. Relative to families that have all members living close to each other, geographic distance makes interactions less frequent and raises barriers to information, both of which can affect the availability and/or costs of commitment and monitoring devices.[3]

With spatial dispersion being a key component of this analysis, I argue that it is important to analyze different dimensions of family behavior for a more comprehensive understanding of family decision-making. In this chapter, I first examine budget allocation decisions. Budget shares have been extensively used in the literature and are widely accepted as a marker of consumption and, therefore, well-being. However, we know that cooperation is more difficult to sustain when outcomes are not easily observable, and one could argue that observing expenditure allocations across households is very difficult, particularly if households are located far away from each other. Additionally, if coordination requires effort, family members might be more willing to coordinate over the subset of goods they really care about and not necessarily over all their allocation decisions. With those ideas in mind, I also analyze family behavior when it comes to investing in the future generations. That is, I look at family investment in child human capital, as indicated by markers of nutrition, education and

---

[3] For evidence on the impact of information on the intra-household decision problem, see Ashraf (2009), Ashraf et al. (2012), Castilla (2010), Castilla and Walker (2013).

cognition. Relative to household budget shares, I understand child outcomes to be very salient to family members and arguably more easily observable. Hence it's possible that family behavior differs across these two dimensions.[4]

To model the family decision problem, I borrow from the methodology developed in the intra-household literature; in particular, I adapt the collective model of the household developed by Chiappori and co-authors and use its testable implications to analyze family behavior.[5] In the context of this theoretical framework, I evaluate three questions regarding the effect of family resources on household outcomes, where resources are measured by the log of per-capita expenditures. First, I look at the extent to which families share resources across households. Specifically, I study whether family resources in the hands of non-co-resident family members affect household outcomes. Next, I test whether the impact of family resources on household outcomes is the same regardless of whether those resources originate inside or outside the household, i.e. the unitary model of the family. Finally, I test whether the allocation of resources is consistent with Pareto efficiency, i.e. whether families are able to achieve cooperation.[6]

---

[4] See Chen (2006) and Chen (2013) for evidence on differences in household behavior based on the observability of the outcomes.

[5] See Chiappori (1988), Chiappori (1992), Browning and Chiappori (1998), Bourguignon et al. (2009).

[6] A few papers have followed a similar methodological approach to analyze families in the United States and Indonesia. For U.S. families, see Altonji et al. (1992), Hayashi et al. (1996), and Dalton et al. (2012), all of which look at household expenditures. Altonji et al. (1992) reject the unitary model of the family, Hayashi et al. (1996) find evidence of risk sharing within families; and Dalton et al. (2012) reject Pareto efficiency. For Indonesian families, see Witoelar (2013), who look at household expenditure; and LaFave and Thomas (2013) who look at child human capital indicators. Both reject the unitary model of the family; Witoelar

Mexico is an ideal setting for this analysis, due to its high rates of internal and international migration, and the relative importance of informal institutions. Mexican migration to the United States is particularly relevant, with a substantial share of Mexican families having relatives living in the U.S. It is very hard to estimate the number of foreign-born individuals living in the United States, but estimates suggest that one-in-ten Mexican-born people live in the United States (Pew Hispanic Center, 2013). Furthermore, evidence suggests that migrants to the U.S. remain connected to their family members in Mexico and remittances are substantial. According to World Bank 2011 estimates, Mexico is the third largest recipient of remittance income across the globe, behind China and India, and the United States is the source of virtually all international remittances.

This research is very demanding in terms of data. We need extensive information on non-co-resident family members and variation in the spatial dispersion among them, including families with individuals spread across international borders. In general, important information such as expenditures and biomarkers is available only for one side of the family, i.e. for the migrant side or for those individuals who stayed in the original location. In this analysis I circumvent these constraints by using a unique multi-

(2013) find evidence of risk sharing within families and LaFave and Thomas (2013) fail to reject Pareto efficiency.

purpose longitudinal survey, the Mexican Family Life Survey (MxFLS), which was explicitly designed to answer questions of this kind.

MxFLS is an extremely rich longitudinal survey of individuals, households, families and communities. The sample is representative of the Mexican population living in Mexico at baseline, in 2002. By design, all individuals interviewed at baseline, as well as children of these individuals born afterwards, are sought for interview in every follow-up, including individuals who move to the United States. Following movers across international borders is a distinctive feature of this dataset and a fundamental element in this setting given the prevalence of international migration from Mexico to the United States.[7] The U.S. component of the survey includes a comprehensive set of modules that closely follow those applied in Mexico, but it also incorporates specific changes to capture important aspects in the lives of Mexicans living in the United States, including information on remittances, use of English, cross-border trips, among others.

At baseline, the sample consists of 35,677 individuals living in 8,440 households spread across 150 Mexican communities. The second wave was implemented in 2005-2007, reaching a 90% re-contact rate. Crucially, this high re-contact rate is not driven by individuals who remained in Mexico. Among individuals thought to be in the U.S. at the time, the re-contact rate was 91%. The third wave spans 2009-2013, and re-contact rates

---

[7] Virtually all international migration from Mexico is to the United States.

are 86% for the whole sample, and 88% conditional on being in the U.S.[8] Using these data, I identify family members through household split-offs, i.e. individuals related by blood or marriage who were living together in a previous wave but are part of different households in later waves. As a result, family in this project is defined to be the group of individuals who were co-resident (part of the same household) in any of the three waves of our panel survey.[9] Finally, through split-offs and migration between waves, I identify different family structures based on the spatial dispersion across family members. In particular, I stratify the sample of families into three groups. First, there is the sample of families that have all of their members living in the same locality in Mexico. Then, there is the sample of families that have members spread across different localities within Mexico. Last, there is the sample of families that have members spread across Mexico and the United States.[10]

Two sets of results emerge from this analysis, depending on whether I analyze household budget shares or child human capital indicators. When I analyze household budget shares, I find that resources in the hands of non-co-resident family members do influence household allocations, but their effect is statistically different from that of

---

[8] These are not final re-contact rates as intensive tracking on the field is not over.

[9] Because I identifying family members through household split-offs, I do not have a complete sample of the family. Having incomplete families is a common limitation in the literature because surveys are not designed to interview families. In Section 2.4, I comment on the implications of having incomplete families and provide some evidence on the degree of incompleteness in these data.

[10] Out of the 1,117 (1,959) families with at least two households in MxFLS2 (MxFLS3), 563 (1,047) have all households located in the same locality in Mexico, 137 (390) families have households spread across Mexico, and 417 (522) families have their households spread across Mexico and the U.S.

household resources. In other words, I reject the unitary model of the family. With respect to whether the allocation of resources is consistent with Pareto efficiency, the conclusion depends on how we interpret the results. While I fail to reject that families are Pareto efficient when I run a joint test that the conditions implied by the model hold, looking at each condition individually I find that more cases than what we would expect to see at random fail to hold. Interestingly, all three results hold regardless of the degree of dispersion among family members.

In contrast, when I look at child human capital indicators, the results differ by family type. For the subsample of families that have all households in Mexico, family resources from outside the household also influence child outcomes. Furthermore, I fail to reject the unitary model of the family, i.e. the marginal effect of family resources is the same regardless of whether those resources originate inside or outside the child's household. However, when I look at international families, resources from outside the household do not affect child outcomes once I control for household resources. Therefore, these results suggest that the *combination* of looking at different degrees of spatial dispersion and different dimensions of family behavior is crucial to a precise understanding of inter-household decision-making.

The remaining of the chapter is organized as follows: Section 2.2 provides additional context on how this work fits into the existing literature, Section 2.3 specifies the model, Section 2.4 describes the data and shows some descriptive analysis of

Mexican families, Section 2.5 presents the empirical specification, Sections 2.6 and 2.7 present the results on budget shares and child human capital indicators respectively, Section 2.8 discusses robustness checks, and Section 2.9 concludes.

2.2 Background

To the best of my knowledge, this work is the first to study family behavior with various degrees of spatial dispersion in a unified framework, including families with members spread across international borders. Existing literature on inter-household interactions in development settings has largely studied either migrant families or related households living close to each other, usually in the same village.[11]

The first line of research corresponds to a large and growing literature whose primary focus has been the study of the determinants and impact of remittances. Virtually all work is based on information collected on one side of the relationship, that is, either from the migrant's point of view or from the family that stayed behind. A unique and important dimension of this analysis is the fact that I look at outcomes and resources from households in both the sending and receiving country.

[11] A third line of research has tried to identify the motives for inter-generational transfers, mainly in the context of developed countries. See for example, Altonji et al. (1997), Bianchi et al. (2007), Cox (1987).

At the other extreme, there is a rich literature on risk sharing within families, but mostly in the context of rural villages.[12] This literature suggests that families do engage in risk sharing agreements and consumption-smoothing activities. However, it seems that these benefits are only evident among households with relatives living in the same village. For example, within poor rural villages in Mexico, Angelucci et al. (2009) show that family members share risk with each other but not with non-relatives in the village. Relative to isolated households, i.e. households without family members in the village, households with relatives in the village invest more when hit by a positive shock and disinvest less when hit by a negative shock. Similarly, Kinnan and Townsend (2012) show that in rural Thailand access to kin networks in the village facilitates access to financing, reduces the sensitivity of investment to income, and helps smooth consumption. Finally, Chiappori et al. (2011) show in the same setting how full risk sharing cannot be rejected among households with relatives in the village, but it can be rejected for households without kin in the village. A common characteristic among these papers is that there is no information on family networks outside the village.

In this study I look at both of these extreme types of families, i.e. international families and families with all members in the same location, where the information on non-co-resident family members is not directly related to their geographic location.

---

[12] See for example, Angelucci et al. (2009), Chiappori et al. (2011), Kinnan and Townsend (2012). For risk sharing within units other than the family, see for example Townsend 1994 (villages) and Munshi and Rosenzweig (2009) (jati).

Linking family members within longitudinal surveys through household split-offs was first proposed by Altonji et al. (1992), and later used by a number of papers, including this chapter. Analyzing food expenditure in PSID, Altonji et al. (1992) look at families in the U.S. and show that, while the unitary model of the family is rejected, resources in the hands of non-co-resident family members influence household expenditures. Dalton et al. (2012) use the same data, and taking advantage of the more detailed expenditure module implemented in later waves, they conclude that resources from outside the family influence housing and education expenditures. Similarly, using the Indonesian Family Life Survey, Witoelar (2013) rejects the unitary model of the family when looking at total expenditures, and finds evidence of risk sharing within Indonesian families. Finally, drawing on the MxFLS, Hamoudi (2010) presents evidence that Mexican families also share risk across households.

Two papers implement the collective model developed in the intra-household literature by Chiappori and coauthors to model the family decision problem, and use the testable implications of the model to test whether the family allocation of resources is consistent with Pareto efficiency. Analyzing household budget shares, Dalton et al. (2012) reject Pareto efficiency for families in the U.S. Looking at human capital indicators, LaFave and Thomas (2013) fail to reject that Indonesian families are Pareto efficient.

As mentioned in the introduction, I follow a similar methodological approach, but none of these papers explicitly looks at the spatial dispersion among family members, including international families. Once adding spatial dispersion into the analysis, I argue that it is important to analyze family behavior along different dimensions. I use two sets of outcomes widely used in the literature: household budget shares and child human capital indicators.

2.3 Model

Family is broadly defined as a group of kin-related individuals who are presumably linked by strong emotional and social ties. This concept would include what is traditionally defined as the nuclear family, that is parents and children, as well as near relatives, such as grandparents, adult siblings, aunts, uncles, cousins and in-laws.

In line with the literature that that looks at interactions across households, this work does not attempt to model either household formation or their geographic location. These are two very interesting processes but are beyond the scope of this work. Likewise, modeling intra-household allocations is beyond the scope of this research.[13]

---

[13] We can still think of the household as a group of individuals with individual preferences, and subsequently household outcomes as the result of some decision process among household members. For our analysis to hold we need to assume that, for the marginal changes we evaluate, the distribution of power within the household is held constant. Formally, we can think of household preferences as defined by: $\mathrm{U}^h(\mathrm{q}, \alpha) = \max_q\{\alpha^1 u^1(q^1) + \cdots + \alpha^N u^N(q^N)\}$ (*Collective household utility function* defined in Browning et al. (2011), where q is a vector of consumption goods, u^k is the utility function of individual *k*, *N* is the total number of individuals in household *h*, and α^k is the weight assigned to individual *k*, weight that is correlated with her bargaining power. Since α will be assumed to be constant throughout the analysis, we simplify the household utility function to be $\mathrm{U}^h(\mathrm{q})$.

Let W represent family welfare, a function of the utility U^h of each household h=1, …,H, with H the number of households in the family.[14] Let q be the vector of private and household-level public consumption goods, q_i^h denote consumption of good i by household h. Let Q be a vector of markers of well-being, such as human capital outcomes of family members, including children. These elements are not goods directly purchased in the market, but produced by the household with a given technology f, vector of inputs d, individual and family endowments σ, and technological parameters δ. Let p^h be the vector of prices of both consumption goods and inputs, expressed in location-1 units, and a and ϵ denote vectors of observable and unobservable preference factors. Assume for the moment that income is exogenous, and let y_h be income from household h expressed in location-1 units, and $Y = \sum_{h=1}^{H} y_h$ denote total family resources.[15]

The family solves the following problem:

$$\max_{v=(q,d)} W(U^1(q^1,\dots,q^H,Q^1,\dots,Q^H;a,\epsilon),\dots,U^H(q^1,\dots,q^H,Q^1,\dots,Q^H;a,\epsilon);a,\epsilon) \tag{2.1}$$

subject to:

$$\sum_h p^h . v^h \leq \sum_h y_h = Y \tag{2.1a}$$

---

[14] I adapt the collective model developed in the intra-household literature. See for example Chiappori (1988), Chiappori (1992), Browning and Chiappori (1998), Browning et al. (1994), Bourguignon et al. (2009).

[15] This model is static. Savings could be thought of as spending on an investment good. Returns on assets can be added as another income source. Moving to a dynamic model is far from trivial, especially in a world without perfect commitment where Pareto weights might change over time (Mazzocco, 2007). I also abstract from labor supply decisions, and assume they are exogenous at the inter-household level.

$$Q_j^h = f_j^h(d_j^h; \sigma, \delta) \tag{2.1b}$$

where f_j^h is a household specific technology for outcome *j*. Solving this problem would lead to reduced-form demand functions of consumption goods *q* and human capital outcomes *Q*. Denote the vector including both by φ.

The model includes household-specific prices to incorporate the idea that the geographic location of the household and its extended family determines the set of prices available to each household. With transport costs added when applicable, I think of the household as facing an envelope of prices among those faced by family members. It is natural to think that families arbitrage and take advantage of geographic differences. While non-tradable goods cannot be consumed at different places, differences in their prices will be connected to differences in wages, and the migration decision was in itself a way to access those wages. Once location decisions are set household-specific prices are taken as given.[16]

If family behavior was consistent with the Unitary model of the family, household demands could be derived straight from problem (2.1), and it would be evident that the distribution of resources within the family is irrelevant. That is, only

---

[16] Note that the main difference with the intra-household literature is not that prices are endogenous to location and past decisions. The price set faced by households is also determined by location and past investment decisions such as schooling. The only difference is that all household members face the same price set, while here prices can vary within families. For that reason, in equilibrium, marginal rates of substitution (which are set equal to prices) will not be the same across households within the family.

total family resources matter. Under such conditions, household demand would be of the form:

(2.2) $\varphi^h = g^h(p^h, Y; a, \epsilon, \sigma, \delta)$

And the testable implication would be:

(2.3) $\frac{\partial \varphi_i^h}{\partial y_k} = \frac{\partial \varphi_i^h}{\partial y_l}$

In words, marginal income effects should be the same across income sources.

The Unitary model of the family is quite restrictive. Therefore, next I evaluate whether the family decision process leads to Pareto efficient allocations, even if resources generated inside or outside the household have different effects on household outcomes. By definition, a Pareto efficient allocation is one such that there is no other feasible allocation that strictly improves the wellbeing of at least one household without making any other household worse-off. Following Chiappori and co-authors, under the assumption that allocations across households within families are Pareto efficient, solution to problem (2.1) should also be the solution to:

(2.4) $\max_{v^1, \dots v^H} \quad U^1(\varphi^1, \dots, \varphi^H; a, \epsilon)$

subject to:

$$\textstyle\sum_h p^h \,.\, v^h \leq \sum_h y_h = Y \qquad (2.4a)$$

$$U^h(\varphi^1, \dots, \varphi^H; a, \epsilon) \leq \overline{U}^{\,h} \;,\; h = 2, \dots H \qquad (2.4b)$$

$$Q_j^h = f_j^h(d_j^h; \sigma, \delta) \qquad (2.4c)$$

for some pre-specified utility level $\overline{U}^{h}$ for each household h= 2, ... H, level that will depend on the weight W gives to each household's utility.

Alternatively, there is a one-to-one relationship between the Lagrange multipliers on constraints (2.4b), and μ^h in the following representation of the problem:[17]

(2.5) $$\max_{v^1,\ldots,v^H} \quad \mu^1\, U^1(\varphi^1,\ldots,\varphi^H; a,\epsilon) + \cdots + \mu^H\, U^H(\varphi^1,\ldots,\varphi^H; a,\epsilon)$$

subject to:

$$\textstyle\sum_h p^h \,.\, v^h \leq \sum_h y_h = Y \qquad (2.5a)$$

$$Q_j^h = f_j^h(d_j^h; \sigma, \delta) \qquad (2.5b)$$

$$\textstyle\sum_{h=1}^{H} \mu^h = 1 \qquad (2.5c)$$

where $\mu^h(p, y, z, a, \epsilon)$ represents the Pareto weight attached to household h, which is a function of prices $p = (p^1, \ldots, p^H)$, income $y = (y_1, \ldots, y_H)$, preference factors (a, ϵ), and distribution factors z, these last ones defined to be variables that do not affect preferences nor do they affect the budget constraint but have an impact on family demand through their effect on the distribution of power within the family.

Up to this point I have allowed household preferences U^h to exhibit unrestricted forms of altruism, and preferences of a given household can directly depend on the consumption levels of other households within the family. If we are going to assume Pareto efficient allocations, under this general form of altruism household allocation

---

[17] Programs (2.4) and (2.5) are equivalent if utility functions U are strictly concave. Furthermore, under strict concavity of preferences one can assume without loss of generality that W in program (2.1) is indeed linear.

decisions cannot be made independently of other family allocations. Therefore, I impose some structure on the form of altruism and assume "caring preferences":

(2.6) $U^h = F^h(\omega^1(\varphi^1; a, \epsilon), \ldots, \omega^H(\varphi^H; a, \epsilon); a, \epsilon)$,

where the consumption of non-co-resident family members enters household welfare only through the "felicity function" ω.

Once preferences are assumed to be caring (or egoistic: $U^h(\varphi^1; a, \epsilon)$), Chiappori (1992) and Bourguignon et al. (2009) show how problem (2.3) can alternatively be solved in two stages. In the first stage, the family agrees on a *sharing rule* θ that assigns to each household a share of total family resources. The distribution of resources will be a function of the distribution of power across households, which in turn is reflected in the Pareto weight μ^h assigned to each household in problem (2.3). In the second stage, each household takes the resources given to it and solves the following problem:

(2.7) $\max\limits_{v^h} \quad \omega^h(\varphi^h; a, \epsilon)$

subject to:

$$p^h . \, v^h \leq y_h^* = \theta_h(p, y, z, a, \epsilon)Y \qquad (2.7a)$$

$$Q_j^h = f_j^h(d_j^h; \sigma, \delta) \qquad (2.7b)$$

$$\theta_h \in (0,1) \qquad \textstyle\sum_h \theta_h = 1 \qquad (2.7c)$$

Denote the solution to these 'individual problems' with $\varphi^{h*} = g^h(p^h, y_h^*; a, \epsilon)$.

From this expression it is clear to see that:

(2.8) $\frac{\partial\varphi_i^{h*}/\partial y_k}{\partial\varphi_i^{h*}/\partial y_l} = \frac{\partial g_i^h/\partial y_h^* \, . \, \partial y_h^*/\partial y_k}{\partial g_i^h/y_h^* \, . \, \partial y_h^*/\partial y_l} = \frac{\partial\varphi_j^{h*}/\partial y_k}{\partial\varphi_j^{h*}/\partial y_l}$, for any two goods i and j, and any two households k, l in family f. That is, while different sources of income are allowed to have different effects on household demands, they satisfy a very particular restriction: the ratio of marginal income effects for any two sources of income is the same across all goods (the ratio is independent of good i). Condition (2.8) will be used to test whether family allocations are consistent with PE.

It becomes clear now why only caring preferences are possible if the model is going to assume Pareto efficient allocations. If more general forms of altruism were allowed in U^h it would not be possible for different households to solve their individual problems (2.7) without taking into account the consumption decisions of non-co-resident family members. Allowing for strategic behavior would lead to inefficient outcomes.

In a more general model we could incorporate family production, family-level public goods and durable goods. If that was the case, the family would decide on the production plan in the first stage, and total income would be allocated to savings, consumption of public goods and durable goods, and the remaining resources assigned to each household.[18]

[18] We should note that whenever we allow for family-level production, family-level public goods, or caring preferences, the existence of a sharing rule θ is only a necessary condition for efficiency. Conditional on the production plan, consumption of public good, and distribution of income across households, household demands can be consistent with constrained Pareto efficiency. However, the overall allocation of resources to production, consumption of public goods and household consumption may not be efficient.

2.4 Data and descriptive analysis

The data used in this project are the Mexican Family Life Survey (MxFLS), an ongoing longitudinal survey that collects a rich set of information on individuals, households, families and communities. The first wave, conducted in 2002, includes 35,677 individuals in 8,440 households spread out across 150 Mexican communities. At baseline, the sample is representative at the national, rural-urban and regional level. The second wave of the survey (MxFLS2) was implemented in 2005-2006, reaching a 90% re-contact rate. The third wave (MxFLS3) spans over 2009-2013 with an 86% re-contact rate.[19]

As briefly mentioned in the introduction, I use the panel structure of the survey to identify families in our data. By design, MxFLS tracks every household member interviewed at baseline in 2002, as well as every child of original household members who are born after 2002. From now on I refer to these individuals eligible for tracking as panel members, and the 2002 household as original household or root household. In later rounds, any panel member that is not part of the original household at the time of the follow-up is sought for interview, and if found is interviewed together with her/his new household members as a new household. In this way, following Altonji et al. (1992), I link every split-off household to their root household, and define this group as a family.

---

[19] Preliminary estimate. Intensive tracking of respondents and data entry are still in process.

A distinctive feature of the data that is essential to this analysis is the fact that panel members living in the U.S. at the time of the follow-up are also followed and interviewed in their new household.[20] Many studies collect information on international migrants from other household members, but few large-scale surveys have tried to follow migrants across international borders.

Without much experience with international tracking, the U.S. component of MxFLS2 was an experiment to test whether such endeavor was feasible. Great effort was put into designing tracking strategies that would lead to finding and interviewing all baseline respondents. This effort paid off, as 91% of those believed to be in the U.S. at the time were interviewed. Due to its experimental nature, interviews were done both in person and by phone, and the questionnaire was shorter than the one applied in Mexico. Nevertheless, the U.S. component of the survey still includes a very comprehensive set of modules that follow those applied in Mexico, while incorporating specific changes to capture the relevant aspects of the life of Mexicans in the United States.

Based on the success achieved in MxFLS2, in MxFLS3 we further expanded the U.S. questionnaire and incorporated important modules such as health measurement. In the third wave we interviewed 88% of the panel migrants living in the US.[21] Specific

---

[20] Virtually all Mexican migrants remain in Mexico or go to the United States. In our data, we only have 18 individuals who migrated outside of these two destinations (for example, to Canada, or the United Kingdom).

[21] Preliminary estimate. Intensive tracking of respondents and data entry are still in process.

differences across waves or across Mexico and the U.S. that are relevant for our analysis will be pointed out below.

2.4.1 Data structure

In Table 15 I present some statistics that illustrate the basic structure of the data. The top panel shows the original sample. At baseline there are 8,440 households which constitute our original sample of families from which some members will split over time. Of those original households, a few are no longer eligible for tracking because all household members died. This leaves 8,386 families in MxFLS2 and 8,335 in MxFLS3.

The middle and bottom panels present the number of households and families interviewed in the second and third rounds. In MxFLS2 we interviewed 8,940 households, 28% of which have related households in the data. By MxFLS3 almost half of the 9,928 households interviewed have related households in the data. This leaves a sample of 1,117 families in MxFLS2 and 1,959 families in MxFLS3. In terms of location, 38% of the families in MxFLS2 and 27% of those in MxFLS3 have at least one household interviewed in the U.S.[22]

To explore whether the impact of family resources on household outcomes changes as a function of the geographic dispersion among related households, families are classified in three groups: families that have all households interviewed in the same

[22] Households interviewed in the last phase of intensive tracking are not included in this table. I will also inform on the location of many households that were found but not interviewed (refusals).

locality in Mexico (*neighbor families*), families that have members spread across different localities within Mexico (*spread across Mexico*), and families with households interviewed both in Mexico and the U.S. (*spread across Mexico and the U.S. / international families*).

Out of the 1,117 families there are in MxFLS2, 563 have all their households located in the same locality, 137 have households spread across Mexico, and 417 have households spread across Mexico and the U.S. In MxFLS3 the 1,959 families are divided into 1,047 neighbor families, 390 families spread across Mexico and 522 international families.[23]

2.4.2 Family in the data

As mentioned before, the identification of family members within the survey comes from linking split-offs with their original household. In other words, a family in this analysis is a group of individuals who were part of the same household at baseline, in wave two or in wave three.

Identifying non-co-resident family members in this way brings two points to discuss. The first one has to do with describing the relationship there is among the individuals we link across households. The second point has to do with the fact that we only observe the fraction of the family that is in the data.

---

[23] There are 8 families in MxFLS2 and 13 families in MxFLS3 that have all their households living in the same region in the U.S., when the U.S. is divided in 5 regions: California, Texas & South, Illinois & Northeast, Midwest & West. Given that I classify families based on whether all members live close to each other or not, these families are included in the second group of families. Considering the very low number of such cases, I still interpret this group as reflecting families split within Mexico. None of the results change if I classify these families as neighbor families or as international families.

With respect to the first point, I checked the relationship that each split-off in the second and third rounds had with the household head at baseline. I find that virtually everyone is a close relative, with adult children and grand-children accounting for the great majority of the cases (almost 80%).[24] Using this information, I identify the sample of families that are connected across households through a parent-child relationship. About 87% of the families in MxFLS2 and 81% of the families in MxFLS3 have such relationship. If I also incorporate the most immediate family, i.e. siblings and grandparents/grandchildren, then 97% of MxFLS2 families and 87% of MxFLS3 families fall in that category. The remaining families are linked through individuals that were mostly cousins, aunt/uncle, and nephew/niece. In the analysis I use all these families, regardless of the relationship there is across households. However, we might think that the parent-child relationship is particularly different than the rest. Therefore, I also run all the models on the sub-sample of families that have this relationship. I find that all the results presented in the following sections remain.

A related point worth discussing, that is not directly related to the relationship across non-co-resident family members but is nonetheless relevant to our definition of family, has to do with the definition of household that we use in the United States. The reason is that it is not uncommon to find migrants living together with the only purpose

[24] I eliminate from the analysis the few split-offs that were not related to the 2002 household head (9 households in MxFLS2 and 14 households in MxFLS3). Results do not change if I keep those observations.

of saving on rent (and utilities), but who otherwise have nothing in common. Therefore, if we had followed in this setting the standard definition of household applied in Mexico most likely we would have not identified the relevant unit of interest for the analysis.[25]

To identify household members in MxFLS2, the questionnaire asks first for the number of individuals living in the dwelling and then for the number of individuals who share income. As it turns out, 86% of households share their dwelling with individuals with whom they do not share income. During the pilots leading to MxFLS3 we explored different definitions and decided on one that emphasized the sharing of expenses besides housing and food. More explicitly, a household was defined as "*a group of individuals who usually live together, usually consume meals provided by a common budget and usually share other expenses*" (besides housing and food). The data show that almost 60% of the households share their dwelling with non-household members as defined above. Once we identify household members in this way, the resulting household structure is very similar to that in Mexico.[26]

Coming back to the second point regarding the incomplete measure of the family in the data, it is unfortunately a prevalent limitation within the extended family literature. Furthermore, it is a limitation that will not be completely overcome until we

[25] Household in the questionnaire applied in Mexico is defined as "*a group of individuals living together and eating from a common pot*".

[26] For more details on the household structure of the U.S. and Mexican samples, as well as differences across MxFLS2 and MxFLS3 see Table 25. Table 26 presents some summary statistics on household and family socio-demographic characteristics, differentiating across family groups, as well as within family groups for those who have members spread across different locations.

implement surveys that are designed to sample families. Families identified from longitudinal household surveys will only be complete families if the survey is extremely long, and just as important, if it has very low levels of attrition. That is very expensive and extremely hard to execute, especially in a setting with high mobility rates, which is necessary to explore differences across families with various degrees of spatial dispersion among its members. While this analysis is based on a relatively short panel, we put a lot of resources into following panel respondents across time and space, including those who move across international borders.

Without a random sample of households within the family, it is very hard to sign the expected bias. On one hand, it could be that we are biased towards cooperation if our sample has those family members who are more likely to be connected, as they lived together at some point in the last 10 years. On the other hand, it could be that interactions across households are more important once those who split are more established, which are more likely to be the links we are missing.

Nevertheless, I can inform on the degree of incompleteness of families in my sample with an unusual level of detail. In the first place, I can identify the number of immediate-family missing links from the report that each adult in the household makes about each non-co-resident sibling, children or parent. Moreover, each adult is also asked whether he/she received help from or provided help to each of these relatives in the last 12 months, differentiating between help in the form of monetary transfers, goods, or time. I also have an overall report on help provided to or received from other individuals. I will use this information to complement the number of missing links with their potential importance as measured by the self-reported intensity of interactions in the last year. In an attempt to address the incomplete nature of the family I will test whether the results are driven by families with more or less missing links in my sample.[27]

2.4.3 Interactions among non-co-resident family members

Even though families can operate under a common agreement without any specific transaction taking place in the time window covered by the survey, such as a monetary transfer, it is useful to have an idea of how much interaction there is among family members. In the data we ask every adult member about help provided to and/or received from each non-co-resident family member, differentiating among children, parents, siblings, and other relatives. We collect information on transfers, which can be in cash or in kind, and for the subsample interviewed in Mexico we also ask about the time individuals spent assisting other family members (e.g. taking care of children). Using these individual reports I construct household-level statistics on the incidence and

[27] LaFave and Thomas (2013) attempt to address this issue by showing how results are robust to a couple of sample stratifications. First they note that results do not change when they split the sample into those whose root household is on the maternal side of the family versus those whose root household is on the paternal side. Then they divide the sample into families that have the majority of individuals in the sample and families for which most family links are missing, and conclusions do not change either.

magnitude of help provided across non-co-resident family members.[28] Table 17 presents some summary statistics, differentiating by family group as well as by location within families. For families spread across Mexico, I differentiate between those interviewed in the baseline location (in 2002), and those interviewed somewhere else. For families across the border I differentiate between those interviewed in Mexico and those interviewed in the U.S. All statistics correspond to the 12 months previous to the interview date.

The first thing to note is that in over half of the households there was at least one individual who either helped or received help from outside the household. The share of households who did not interact with non-co-resident family members is 45% among neighbor families, 44% among families spread across Mexico, and 29% among international families.

A second observation is that the sample of neighbor families and the sample of families spread across Mexico look very similar to each other, as do households interviewed in different locations within the latter group of families.[29] In contrast, international families are, on average, very different from the rest. Furthermore, within these families, household interviewed in Mexico and those interviewed in the U.S.

---

[28] Statistics regarding to whom respondents provided help, e.g. parents, remain at the individual level.

[29] The only difference within families spread across Mexico is that individuals interviewed in a new location are more likely to have sent transfers to non-co-resident siblings relative to individuals who remained in the original location (both conditional on having siblings outside the household).

present very different statistics regarding incidence, direction and magnitude of transfers.

With respect to the two samples of families with all households in Mexico, about 26% reported both helping and receiving help from non-co-resident family members, 18% only provided some help, and 12% only received some help. In terms of how they helped each other, households are more likely to transfer money or gifts rather than dedicating time to assist others (about 40% and 10% of households respectively). Finally, the magnitude of transfers sent and received are pretty similar. On average, households sent about US$2,000 in the last 12 months to non-co-resident family members (US$4,500 conditional on having sent some), and received about US$1,700 from non-co-resident family members (US$4,500 conditional on having received some). These magnitudes represent between 40% and 50% of total household expenditures. If we compute net transfers by subtracting the amount received from the amount sent, households sent on average about US$500 to non-co-resident members (conditional on either sending or receiving some help).

With respect to international families, the distribution of those who only provided help, only received help, or both, is substantially different between households across the border. In Mexico, 29% reported providing and receiving help from non-co-resident family members, 21% only received help and 15% only provided help. In contrast, in the U.S. only 4% reported only receiving some help while 57% reported only providing some help (and 17% did both). Similarly, the net help received is significantly different across the border. On average, households in Mexico are net receivers. They received, in the 12 months prior to the interview, about US$3,000 worth of transfers. In contrast, households in the U.S. sent on average US$3,000 in the same time period.

Table 28 provides additional summary statistics on international remittances sent by Mexican migrants in the U.S. In terms of the intended use of the transfers, we see that conditional on having sent some, the majority of the migrants sent remittances for the final consumption or use of the recipient, and only a few reported having sent some amount for saving or investment purposes (92% versus 14%). We also ask each migrant to identify the beneficiaries of the transfers. Using that information we get that 40% of the migrants reported having sent transfers to more than one individual in Mexico, with roughly half of them helping two individuals and half of them helping three or more individuals.

Finally, we collect detailed information on the magnitude, means and cost of the last transaction. Using information on the total value of transfers sent in the last 12 months, and the amount sent in the last transaction, we can infer the frequency with which migrants sent transfers home. Assuming that the last transaction represents a regular transaction, as suggested by the testimony we got while doing the pilot interviews, migrants make on average 17 transfers a year (median of 10). Most of these transactions are made through a financial institution, with only 7% of the migrants using relatives, friends or themselves to take the money to Mexico.[30] Finally, we asked about the cost paid to make the transaction. The reported distribution has a high density at two values: 10% reported 0, and 63% reported US$10. On average, these costs represent 6.5% of the value of the transfer (median of 5%).

2.4.4 Outcomes of interest

As mentioned in the introduction, I estimate the model with respect to two sets of outcomes: household expenditure shares and child human capital indicators. Both sets of outcomes have been widely analyzed in the literature to test the unitary model of the household or to test whether households are Pareto efficient. At the family level, household expenditures have been mostly used to test risk sharing within families and to test whether families are altruistic, which is empirically equivalent to testing whether families are unitary. To the best of my knowledge, only LaFave and Thomas (2013) use child human capital outcomes to test Pareto efficiency at the family level.

In this chapter I look at both sets of outcomes, and consider adding more in the future, to have a more comprehensive view of family behavior. This is particularly important as we explore whether conclusions change when we look at different types of families defined by the spatial dispersion among its members. Family members that live far away from each other have fewer opportunities to observe each other's behavior. As

[30] We collect information on both monetary and in kind transfers. On average, 17% of the value transferred to Mexico was in kind.

a result, families might be more likely to cooperate on outcomes that are more easily observable, or family members might be more willing to cooperate over outcomes that are really important to them.

With this in mind I first evaluate household expenditures. This is the natural starting point if we want to see to what extent resources from non-co-resident family members affect household resource allocation decisions. It is an outcome that has been widely used in the literature, and taken as a marker of consumption it is a determinant of individual and household welfare. Additionally, our data is unique in collecting expenditure information on both non-migrant and migrant households.

However, one could argue that, for the most part, expenditures are difficult to observe, and are likely to be measured with error. Additionally, family members might not be as concerned about expenditures on certain goods such as clothing or entertainment activities, as about child wellbeing. Therefore, I also study the family resource allocation process when it comes to investing in the next generation. By looking at child human capital indicators, I evaluate family behavior over outcomes that are presumably of great interest to family members, and to some extent more easily observable.

2.4.4.1 Budget shares

To have expenditure data on households located in Mexico as well as on their relatives living in the United States is a very unique feature of these data. In Mexico, the expenditure module is the same across the three waves. In the U.S., due to the experimental nature of MxFLS2, the second wave only collects information on total household expenditures. Therefore, the sample of international families in the second wave only includes households interviewed in Mexico. In MxFLS3 we extended the questionnaire and designed the expenditure module following closely the one applied in Mexico. As a result, I can define the same bundles of goods and construct comparable expenditure shares across the two countries.[31] About 6% of households in the data have missing expenditures.

To do the analysis I need to divide total expenditure in groups of goods, taking into account the trade-offs between aggregating over goods that are very different on the one hand, and having too many observations with zero expenditure on the other. High density at zero introduces a selection bias that very hard to correct for, as there aren't good instruments that correct for the two underlying sources of zero expenditure, i.e.

[31] The modules across the two countries are not exactly the same. Two main differences are the level of disaggregation in the questions (e.g. all food versus types of food), and for a few items (education and semi-durables) the reference period, which in the US does not go beyond 3 months prior to the interview date. In terms of the level of disaggregation, when designing the questionnaire we faced the trade-off between comparability with the module in Mexico and the length of the interview in the US. Since it was the first experiment with a face-to-face interview, and the questionnaire is quite long, we compromised by selecting broader categories. With respect to the reference period, we did not want to ask for a period too long before the interview date in order to minimize the probability that the reference period covers both time spent in the U.S. and time spent in Mexico for very recent or circular migrants. Even though we expect these two differences to affect reported expenditure, the assumption we need is that expenditure shares are not affected. When we use total expenditure as our measure of resources on the right-hand-side, we need the instrument not to have an heterogeneous effect along the expenditure distribution, so that the difference in levels is appropriately controlled for with a wave dummy.

the household does not consume that good, or the household did not spend on that good in the time window specified in the questionnaire.

Considering both the nature of the goods and the share they represent in total expenditures, I do the analysis on 5 groups of goods.[32,33] These are: food, which includes food consumed at home and meals outside; personal care, clothing, health, education and recreation; transport and communication; semi-durables, cleaning supplies, insurance and car or house repairs; and housing, which includes the rental value of the dwelling and utilities.[34] In Table 17 I present summary statistics for all families and by family type. In Table 27 I further present statistics differentiating within family groups among households interviewed in different locations (i.e. international families are split into those interviewed in Mexico and those interviewed in the U.S.), and test whether differences across families, and within families, are statistically different.

2.4.4.2 Child human capital

I estimate the model on three markers of child development: height-for-age z-scores, which are a marker of nutrition; years of education; and ravens scores, a marker of cognition.

---

[32] All magnitudes are converted to monthly expenditures. Expenditures on durables are excluded from the analysis.

[33] While a complete analysis on budget allocation decisions would include savings, I follow the tradition in the literature and abstract from savings decisions because I do not have a good way to estimate how much households save in a given period.

[34] For the case of home owners, housing includes the self-reported rental value of their dwellings.

With respect to the first outcome, the sample includes children interviewed in Mexico in either wave and children interviewed in the U.S. in the third wave.[35] All health measures are taken by trained personnel, and using the same instruments and following the same protocol in both countries. I standardize height by calculating height-for-age z-scores in order to control for age and gender specific factors. I calculate age in months at the time of measurement based on date of birth and interview date, and follow the 2000 CDC Growth Charts, which uses a representative sample of well-nourished children in the U.S. as the norm.

For this outcome I restrict the sample to children between 1 and 4 years old at the time of measurement. Evidence from the nutritional literature suggests that nutritional inputs affect child height only during the first years of life. While there is some controversy around the exact age cut-off over which height is no longer sensitive to nutritional inputs, there is a consensus that it ranges between the ages of two and four. Based on this evidence, I allow for current household and other family resources to affect child height up to the age of four. Taking into account the relatively small sample

[35] For the same reasons mentioned above, MxFLS2 does not collect anthropometrics and biomarkers for the sample interviewed in the U.S. Note that this does not mean we do not have a sample of children in international families in MxFLS2. As long as we have a measure of resources from U.S. households we can include in our analysis children interviewed in Mexico that have family in the U.S. Furthermore, MxFLS2 households in the U.S. are formed by recent migrants and therefore few of them have young children living in the U.S. with them. The data suggests there are 52 children we could not include due to lack of health measurement in the U.S.

size and the high measurement error involved in measuring very young children I exclude those younger than one at the time of measurement.

With respect to years of schooling and ravens scores, I use children six to sixteen in the first case and children five to sixteen in the second case. Five is the youngest age at which children are given the cognitive test. While I could add five-year-olds to the schooling sample, it does not add value to the analysis as virtually every kid in Mexico attends some sort of kindergarten but formal schooling does not start until the age of six.

The cognitive test consists of a series of Raven's Colored Progressive Matrices (CPM) which aims to assess non-verbal cognitive function. It is considered a measure of general intelligence, and accepted as the single best measure of Spearman's general intelligence factor (Raven, 2000; Kaplan and Saccuzzo, 1997). I take each child's answers and calculate the percent correct. Finally, the sample of children used for this outcome is children interviewed in Mexico, as this module is not part of the U.S. questionnaire in either wave.[36]

In Table 18 are summary statistics on the outcomes, household and parent characteristics for each sample children, differentiating by family group. Differences in mean outcomes within families among those interviewed in different locations, and tests

[36] Contrary to the case of height-for-age, the sample lost due to lack of measurement is considerably larger, as it includes children age 5 to 16 either in MxFLS2 or MxFLS3. To mitigate concerns related to sample selection in a robustness analysis presented in Section 2.8 I restrict the sample to children interviewed in Mexico for the three outcomes and find that results remain the same.

of whether differences between families and within families are statistically significant are presented in Table 27.

2.5 Empirical specification

From the model presented in Section 2.3 we derived a conditional demand function of the form:

(2.9) $q^{h*} = g(p, y_h^*(p, y, z, a, \epsilon, Y); a, \epsilon)$, with $y = (y_1, \dots, y_H)$

To test the empirical implications of the model I estimate the following linearized version of demand functions:

(2.10) $out_{khf}^i = \alpha^i + \tau_1^i(y_h) + \tau_2^i(y_e) + x_{khf}^{'}\gamma^i + \varepsilon_{khf}^i$,

where out_{khf}^i is outcome i of individual k in household h in family f (out_{hf}^i if household-level outcomes), $\tau_1^i(y_h)$ is a flexible function of household resources, $\tau_2^i(y_e)$ is a flexible function of extended family resources (or other family resources, or resources from non-co-resident family members), and $x_{khf}^{'}$ is a vector of individual, household and family characteristics.[37]

I estimate alternative specifications of system (2.10), which differ on the functional form assumed on $\tau\,(.)$. First I estimate a linear demand system. Denoting β_1^i to be the marginal effect of household resources, and β_2^i the marginal effect of other family resources, the tests are:

---

[37] Throughout the paper, *family resources* refers to the sum of household resources for the complete family, while *extended family resources* are total family resources minus resources from within the household.

Unitary test: $\beta_1^i = \beta_2^i$ for every outcome i

Pareto test: $\beta_1^i / \beta_2^i = \beta_1^j / \beta_2^j$ for any pair of outcomes i, j

I estimate a seemingly unrelated regressions demand system (SUR), with cluster standard errors at the family level. To test Pareto efficiency we need to implement cross-equation tests. I estimate non-linear Wald tests calculated using the delta method allowing for clustering at the family level. I re-express the test as the cross-product of coefficients instead of ratio of coefficients. I present the p-value associated with each individual test across any two outcomes, as well as the p-value associated with the joint test that the ratio holds across all pairs of outcomes simultaneously. By showing pair-wise tests we can identify in which cases Pareto efficiency fails, and results are less sensitive to the number of goods included in the system.[38]

I also explore a number of non-linear specifications. Using splines with alternative cut-offs, I allow the marginal effect of household and other family resources to be different at different parts of the distribution. In this case, to test whether the marginal effect of family resources from within and outside the household is the same, i.e. the unitary test, I see whether $\beta_1^i = \beta_2^i$ for each spline, as well as whether it holds simultaneously across the entire distribution. To test for Pareto efficiency, I test whether the ratio of marginal effects is the same across any pair of goods and for each possible

[38] The critical value of the Wald test depends on the number of goods included in the system, and therefore it is possible to fail to reject PE by adding enough goods.

combination regarding the position of household and other family resources in the distribution. For example, if household and other family resources are allowed to have a different effect on budget allocations when they are below and above the median, then I test: whether the ratio of marginal effects between goods *i* and *j* is the same when both household and other family resources are below the median, when both are above the median, when household resources are below but other family resources are above the median, and when household resources are above and other family resources are below the median. I do the same for each possible pair of goods. Once I exhaust all pair-wise combinations, there are a number of subgroups we might be interested in testing simultaneously. In this analysis I present the joint test across all pairs of goods for each possible case within the family, i.e. each combination of the relative position of household and other family resources. At last I present whether all conditions hold simultaneously.

All values are expressed in 2009 U.S. dollars. I use annual PPP exchange rates to convert Mexican pesos to dollars, and I use national inflation rates to express magnitudes in 2009 U.S. dollars. Further differences in price levels are controlled for by location and interview-date variables. All covariates are calculated for each wave (MxFLS2 & MxFLS3) and then observations from both waves are pooled together.

My measure of resources is the log of household per-capita expenditure. Expenditure is widely accepted as a better marker of long-term resources relative to income, which is subject to high variation over the life course, and has less measurement error.[39] While in this particular setting one might worry that household expenditure is already the result of the family decision problem, it is not clear that using income addresses this concern. Families also organize the labor supply of their members, choosing who works and where they work. Therefore, I consider household expenditure to be the best measure of resources at hand.

In Table 17 I show summary statistics on household, extended family, and total family per-capita expenditure, for all families and for each family group separately. In Table 27 I further show summary statistics within family groups, differentiating among households interviewed in different locations, and tests of whether differences across families and within families are statistically different.

In the absence of exogenous variation in household and other family resources, I follow the literature and control for a wide set of observable household and family characteristics. Fortunately, these data collects a very extensive set of outcomes, including markers not generally observed in household surveys.

When the outcome of interest are household budget shares the vector of controls include: household size and composition (number of children, prime-age males, senior males, senior females, with prime-age females as the omitted category); age, gender and years of education of household head; household location (dummy of whether the

[39] Additionally, there is a non-trivial share of households in the data with zero income.

household is in the US, region within the U.S., state within Mexico and whether the place is rural if the household is in Mexico); interview date (quarter-year dummies); MxFLS wave (dummy if observation is from MxFLS2, and interaction between U.S. and MxFLS2); and family characteristics (age of family head; maximum years of education, height and cognitive score among adult males; maximum years of education, height and cognitive score among adult women; family size, number of children, prime-age males, senior females and senior males; log of per-capita wealth at baseline in 2002, its square, and its cube).[40]

When evaluating the effect of resources on child human capital indicators the baseline set of controls are: age dummies, sex, whether the household is in rural place, household size and number of kids under the age of 15, State in Mexico, quarter-year interview dummies, family size, number of kids under 15 and number of female adults in family, U.S. dummy, MxFLS2 dummy, and the interaction of U.S and MxFLS2. In Section 2.8.1 I present estimates when I include additional controls in the estimation.

2.6 Results: household budget shares

I start the analysis by adopting the simplest specification, a linear model, and by pooling all families together. Table 18 shows the estimated effect of household and other family resources on household budget allocations across the five main consumption groups. Column 1, for example, suggests that a 1% increase in household per-capita

[40] The United States was divided into: California, Texas & South, Illinois & Northeast, Midwest & West.

expenditures is associated with a 4.13 percentage point decrease in food shares, effect that is highly significant. In contrast, the negative effect of resources in the hands of non-co-resident family members, though statistically significant, is much smaller: a 1% increase in per-capita expenditure outside the household is associated with a 1.3 percentage point decrease in household food shares.

More generally, holding other family resources constant, as household resources increase expenditure shares in food and housing decrease while expenditure shares in other goods increase. Holding household resources constant, as resources in the hands of non-co-resident family members increase, the household budget allocation to food decreases, the allocation to semi-durables and housing increases, and the allocation to personal goods, transport and communications does not change. A joint test across the five outcomes suggests that resources from non-co-resident family members are statistically significant in determining household budget allocations.

At the bottom of Table 18 I present, for each budget share, the difference between the estimated marginal effect of household and other family resources, and the p-values associated with the test that the difference is zero. For every outcome we reject that that is the case. In other words, while resources from outside the household matter, their effect is significantly different from that of household resources, i.e. the family does not behave as a single unit.

To see whether the evidence is consistent with Pareto efficiency, I present the ratio of β_1^i/β_2^i for each outcome, as well as the p-value associated with the test that the ratios are the same across the 5 columns. The p-value associated with each pair-wise test is presented in Table 32. Based on this analysis, Pareto efficiency is rejected.

While the rejection of the unitary model is not specific to this specification, I show next that the rejection of Pareto efficiency is. It is enough to either estimate the system on each family group separately (see Tables 30 and 32), or to allow for a more flexible specification for the results to change.

In Table 19 I present the results of my preferred specification, which allows for the marginal effect of own and other family resources to be different at three levels of resources: below the 25^{th} percentile, between the 25^{th} and 75^{th} percentile, and above the 75^{th} percentile.[41] For easy of interpretation, the cut-offs are based on the distribution of the log of family per-capita expenditure when I pool all families together. Therefore, they are unique across family groups as well as across household and other family resources. Nevertheless, results remain the same if the cut-offs are defined for each family group and each source of resources separately.[42]

---

[41] See Tables 31 and 33 for a specification that allows for resources to have a different effect below and above the median. The results suggest that the conclusions of this section are not unique to my preferred specification.

[42] See Table 17 for summary statistics on the log of household, extended family, and family per-capita expenditure, for the whole sample as well as by family group.

2.6.1 Neighbor families: all households interviewed in the same locality in Mexico

Results are shown on the right panel in Table 19. Just as we saw when pooling all families together, as household resources increase budget allocations to food and housing decrease. However, these effects are very non-linear. For food shares, the strongest negative effect takes place at the middle of the distribution. In the case of housing, the strongest effects are at the bottom and top of the distribution. For both outcomes, a test that these effects are the same along the distribution is rejected (see bottom panel in Table 19). Expenditure shares in the other three consumption groups increase as household resources increase, and we fail to reject the effect is linear, except for the case of transport and communication.

Even after controlling for household resources in a more flexible way, resources of non-co-resident family members have a significant influence on household budget allocations. In particular, they are statistically significant for food, personal goods and housing budget shares. With respect to food shares, higher extended family resources are associated with lower food shares. Even though this non-linear specification does not show significant effects, the hypothesis that the effect is linear cannot be rejected, and the linear model shows a negative and highly significant point estimate of -1.41 (see Table 30). With respect to personal goods (personal care, clothing, health, education and recreation), other family resources are significant at the top of the distribution, increasing the budget allocation towards such goods. Finally, other family resources below the 25th percentile seem to be positively associated with a higher housing share, though we can't reject that this effect is constant across the entire distribution. In a linear specification, the positive effect remains positive and significant (see Table 30).

The bottom panel of the table shows p-values associated with the unitary and Pareto efficiency tests derived before. Overall, the unitary model is rejected for all outcomes. Only in a few cases, when we focus on specific parts of the distribution, I fail to reject the null that the marginal effect of household and other family resources is the same. These are: for food budget shares if resources are very low or very high, and for the expenditure share on personal goods if resources are very high.

As mentioned before, testing for Pareto efficiency in a non-linear model is less straightforward. In Table 20 I present the p-values associated with the test that any two ratios are the same, for each pair of goods and each possible combination regarding the position of household and other family resources. For example, for the case that both household and other family resources lie below the 25th percentile, I test whether the ratio of marginal effects is the same across any two budget shares. Similarly, for the case where household resources lie below the 25th percentile, but other family resources lie above the 75th percentile, I also test whether ratios are the same for each pair of goods. And the same follows for any other combination of resources. At the bottom of the table, I show the joint test that the ratios across all pair of goods are simultaneously equal to each other for each possible resource combination. And finally, I run an overall test across all goods and resource combinations. The p-value of this last test is the one shown in Table 19.

If we look at the bottom of Table 20, we see that I never fail to reject that these families are Pareto efficient, regardless of the relative position in the resource distribution of households and their extended families. However, if we look at each of the individual tests, it seems that more cases than what would be randomly expected are statistically significant. If I count the number of cases that are statistically significant at the 5% level, and weight each case by the share of households that are in the corresponding household/other family resources cell, I find that 7% of the tests are statistically significant.

2.6.2 Families spread across Mexico

Next I look at the effect of family resources when households are spread across different locations in Mexico. With respect to household resources, the general patterns we saw on neighbor families remain, though some point estimates are not statistically significant when resources are above the 75th percentile. In contrast, the effect of family resources from outside the household is less clear than what we find for the other two family groups. Resources in the hands of non-co-resident family members are only statistically significant for two outcomes, food and semi-durables, and only at the bottom of the distribution. When I test whether these effects are statistically significant across the three segments of the distribution, I only reject the null that they are not at the

10% significance level when it comes to food shares (p-value of .07). Finally, I fail to reject that these resources do not matter for household allocations when I test across the five outcomes.[43]

Once again, I reject the unitary model of the family. The only case in which I fail to reject that marginal effects are the same is for personal goods. As to whether allocations are consistent with Pareto efficiency, I fail to reject that they are when I test that all ratios across any two goods are the same, regardless of where household and other family resources lie in the distribution (see bottom panel in Table 20). However, if I look at the individual tests I find that I reject the null in 5.1% of the cases.

2.6.3 International families: households spread across Mexico and the United States

Finally, I look at budget allocations among international families. The general shape of the effect of household resources on budget shares remains similar to the previous cases. As household resources increase, household budget shares on personal goods, semi-durables and transport and communication increase. At low and high levels of household resources, these increases take place at the cost of decreasing housing expenditures. At middle levels, resources are mostly shifted away from food. As resources from non-co-resident family members increase, households allocate more

[43] For this group of families, I fail to reject that the effect of resources from non-co-resident family members is linear. Nevertheless, even if I impose a linear specification these resources are not statistically significant (See Table 30).

towards housing, and allocate less towards food, personal goods, and transport and communication.

Once again, I reject the unitary model of the family. Just as we saw for neighbor families, I only fail to reject it in a few cases if I look at specific parts of the distribution. These cases correspond to situations where neither household nor other family resources have a significant influence on budget shares, which only occurs for some outcomes at very low or very high levels of resources. Finally, at the bottom of Table 20 we see that I always fail to reject that allocations are Pareto efficient whenever I test simultaneously that the ratios are the same across all pairs of outcomes. However, if I count the number of individual tests that turn out significant, I reject the null that ratios are the same in about 6.5% of the cases.

To sum up, the analysis on household budget shares suggests three things. First, resources in the hands of non-co-resident family members influence household allocation decisions (at least for neighbor and international families). Second, changes in household resources and changes in resources from non-co-resident family member have different effects on household allocation decisions. In other words, the unitary model of the family does not hold. Finally, when I test whether the ratio of marginal effects is the same across any pair of goods, I fail to reject that they are, which is consistent with Pareto efficiency. However, if I count the number of individual tests that are significant at the 5% level, there are more cases than what would be expected to see at random. This holds for all three family groups.

2.7 Results: child human capital

Table 21 presents the main results, which for the case of child outcomes correspond to the linear model.[44]

When pooling children from all family groups together, we note that both household and extended family resources are positively associated with the outcomes of interest. After controlling for extended family resources, children in households with higher per-capita expenditures are taller, have more years of education, and score higher in the cognitive test. Similarly, after controlling for household per-capita expenditure, the estimated effect of resources in the hands of non-co-resident family members is always positive and significant. Point estimates suggest that the effect of own resources are stronger than that of non-co-resident family members (the ratio of β_1^i to β_2^i ranges between 1.6 and 1.92). However, when I formally test the unitary model ($\beta_1^i - \beta_2^i$=0), I fail to reject that these two coefficients are the same. Consequently, I also fail to reject Pareto efficiency.

2.7.1 Families in Mexico

The second and third panels in Table 21 suggest that the patterns we just described when pooling all children also hold for the two family groups that have all

[44] For non-linear specifications see Section 2.8.3.

households in Mexico. Among families with households living in the same locality, and among families with households spread across Mexico, family resources seem to have the same impact on child outcomes regardless of whether those resources come from within or outside the child's household. In other words, in both cases I fail to reject the unitary model of the family. Therefore, I also fail to reject Pareto efficiency. Table 22 shows individual tests across each pair of outcomes, as well as the joint test across all combinations.

2.7.2 International families

The story is quite different when we look at children in international families. Household per-capita expenditure continues to be strongly associated with nutrition and educational outcomes. However, after controlling for household per-capita expenditures, expenditure levels of non-co-resident family members do not have a significant effect on child outcomes. Consequently, the unitary model is rejected (with the exception of cognitive scores, where big standard errors make the coefficients indistinguishable from each other).

Considering that households in Mexico and households in the U.S. are very different from each other in terms of resources and socio-demographic characteristics (see Tables 26 and 27), it might be important for this sample of families to explicitly differentiate between these two sides of the family. Therefore, in Table 23 I present two alternative specifications.

First I allow for the effect of household and other family resources to be different based on whether the child is interviewed in Mexico or in the U.S. The results are presented in Panel A. We note that neither the marginal effect of household resources nor the marginal effect of other family resources changes with the location of the child (p-values of the test that the estimated coefficients are the same are shown at the bottom of the table).[45] Similarly, in the second panel of the table I present the estimated effects when I differentiate between resources generated in Mexico and resources generated in the U.S. In this case, only extended family resources generated in Mexico turn out significant for height-for-age.

An alternative concern we might have when working with international families relates to the use of household expenditures as a measure of household resources. The reason is that, if transfers across non-co-resident family members are important, household expenditures might not be as tightly connected to resources generated inside the household as they would be in the absence of such transfers. In Table 16 I show that over half of the households in all family groups either received help or provided help to non-co-resident family members in the 12 months prior to the interview date. However, what distinguishes international families from the rest is the noticeable difference there is in the direction of the transfers within the family. While, on average, net transfers in

---

[45] Note that we can do this exercise only for height-for-age and years of education as we do not have cognitive score for children interviewed in the U.S.

the other two family groups are close to zero, households in international families interviewed in Mexico are net recipients of help while households interviewed in the US are net providers of help.

In an attempt to take these differences into consideration, I present in Table 24 results from three alternative specifications. In the first panel I use income and wealth as instruments for household expenditures. The reasoning behind this exercise is that the variation in household expenditure generated by these two other markers of resources might better reflect the initial relative position of the households within the family. However, these results should be interpreted with caution, because the first stage is not very strong. In particular, two F-statistics fall below 10 (household resources in height-for-age, and other family resources in cognitive scores). Nevertheless, under this specification family resources from outside the family do impact child height. In panels 2 and 3 I continue to use household expenditure, but I explicitly take into account transfers across non-co-resident family members. In one specification I consider net transfers as a household expenditure. In the second specification I only consider transfers provided to non-co-resident family members as expenditure (without subtracting out the help received from outside the household). Table 24 shows that both specifications deliver very similar results, and neither adjustment makes resources from outside the household a significant determinant of child outcomes.

To sum up, I do not find strong evidence that after controlling for household resources, resources from outside the household affect child human capital indicators. Differentiating between children in Mexico and children in the U.S., or taking transfers explicitly into account, do not change this conclusion. Only when I instrument expenditure with income and wealth other resources are significant, but only for height-for-age. In additional robustness checks presented in Section 2.8, I show that these results remain if I explore non-linear specifications (Section 8.2.3, Table 36), or if I extend the analysis to other health outcomes (Section 8.2.4, Table 37).

2.8 Robustness checks: child human capital

2.8.1 Additional controls

The baseline specification is very parsimonious because I want to capture the overall effect of resources from within and outside the household. Therefore, in order to present evidence, to the extent possible, that the conclusions are not driven by unobserved heterogeneity, I present next a few specifications that include a rich set of controls into the model.

Table 34 presents four additional specifications. Panels A, B and C add different sets of household or family covariates, and panel D controls for mother's height.

The first panel present estimates when I control for traditional household head characteristics: age, sex, years of education and marital status. This results in overall smaller point estimates that are less precisely estimated. However, the changes affect primarily the estimated effect of household resources and not so much the estimated effect of other family resources. Two point estimates on the log of household per-capita expenditures decrease substantially and are no longer significant: the effect of household resources on years of education for neighbor families, and the effect on cognitive scores for international families. Conclusions on whether the unitary model of the family holds do not change, although the rejection of the model for international families now comes from rejecting the test on one, instead of two of the three outcomes.

Panel B presents the estimated effects when I add to the model relatively less traditional characteristics. These are: household head's height, cognitive score, and risk preferences.[46] All these characteristics should be correlated with a number of unobserved factors that correlate with resources and the outcomes of interest. Nonetheless, we note that estimated effects barely change. The only two differences relative to the baseline specification are: the effect of extended family resources on height-for-age when we look at families spread across Mexico and the effect of household resources on cognitive scores among international families. We still fail to reject the unitary model of the family for neighbor families and families spread across Mexico, and reject the model among international families.

---

[46] MxFLS has a preference module with a set of questions related to time and risk preferences. To estimate risk preferences the respondent is presented with a set of questions asking to choose among different lotteries with various levels of risk. From that module I construct an index that identified relatively more risk averse individuals. This module was not asked to U.S. respondents in MxFLS2.

So far I have added household characteristics, but unobserved factors across families could also affect the results. In Panel C I present the estimated effects when I add a set of family characteristics measured at baseline. These are: the log of household per-capita expenditure, and age and years of education of the household head. The changes in point estimates relative to the baseline specification are similar to those in Panel A. Additionally, results are virtually the same if I add the whole set of family characteristics included in the analysis on household budget shares: detailed family composition at the time of interview; highest education, height and cognitive score among male and female adults; and a polynomial on wealth in 2002 (Results available upon request).

Finally, we know that height has a strong genetic component and the literature has provided substantial evidence on the correlation between attained adult height and earnings. Therefore, I also present estimates that control for mother's height in Panel D. We note that the estimated effects of both household and extended family resources decrease when we look at children's height-for-age, though the coefficients on extended family resources are relatively less affected. This specification also decreases the effect of household resources on cognitive scores for international families, effect that is no longer statistically significant. In spite of these changes, the general patterns emerging from our baseline specification remain.

2.8.2 Alternative samples

In Table 35 I present the estimated effects when I use different subsamples of children.

First I address the concern relating the change in instrument between MxFLS2 and MxFLS3. Panel A presents the estimated effects when I restrict the sample to children interviewed in MxFLS3. Results suggest that, even though I lose precision, the differences across family types identified in the baseline specification remain.

In Panel B I intend to address the issue that we only apply the module on cognitive ability in Mexico. Even though we will never know how results would look like if we had cognitive scores on children interviewed in the U.S., I can look at whether imposing the same sample selection on the other two outcomes changes the estimates. The results at the bottom of the table show that the estimated effects barely changed for either height-for-age or years of education.

2.8.3 Non-linear effects

In this section I present two specifications that show that the conclusions reached previously are not driven by the linear assumptions imposed in the baseline model. The results are shown in Table 36.

In panel A I allow for the effect of resources from non-co-resident family members to have a nonlinear effect on child outcomes. I present estimates that allow for the marginal effect of extended-family resources at the top of the distribution, above the

75th percentile, to be different than the marginal effect below that cut-off. We see that even though the estimated effects are not exactly the same, I fail to reject the null that the effect is linear (p-values shown in the last row).[47] Also important, the lack of significance in the group of international families is not driven by incorrectly imposing linear effects.

In panel B I allow for resources from non-co-resident family members to be different based on whether household resources are low or high. In particular, I estimate the marginal effect of other family resources for households at the bottom half and at the top half of the household per-capita expenditure distribution. The estimated effects are extremely similar across these two groups of households.

2.8.4 Alternative outcomes (international families)

Finally, I evaluate whether the lack of significance in the effect of resources from non-co-resident family members is specific to the three outcomes I evaluate. For example, for the case of cognitive scores, we only have information collected in Mexico. With respect to child height, we know that there is a short window over which resources can affect it, as the nutrition and biological literatures have established that the growth path is pretty much established by age 4.

To see whether our results are robust to analyzing a broader set of outcomes, I estimate the system on a wider set of health markers, all of which are responsive to resources through childhood (and adulthood) and are collected both in Mexico and the

[47] Same results hold if I use the 50th percentile as the cut-off.

U.S. I evaluate BMI-for-age, an indicator of whether BMI is lower or equal to 18.5, weight-for-age, hemoglobin, and an indicator of whether hemoglobin is below 12.[48] I also added an indicator of whether the child is stunted, which is defined as having a height-for-age z-score lower than -2.[49] The results are in Table 37. Note that resources from non-co-resident family members are never significant. This holds both in a linear model and in a model that estimates different marginal effects for the bottom and top half of the distribution. The only exception is on the indicator of low hemoglobin levels, where the coefficient on other family resources is marginally significant.

2.9 Conclusion

This analysis contributes to the literature that looks at interactions among family members when they do not live together, recognizing the important role of the extended family in shaping individual and household outcomes. I first analyze the extent to which Mexican families share resources across households, i.e. whether resources from non-co-resident family members influence household resource allocation decisions. Finding that they do, I then see whether the family behaves as a single unit, i.e. whether the marginal effect of family resources on household outcomes is the same regardless of whether those resources originate inside or outside the household. Finally, I test whether family

---

[48] Both cut-offs on BMI and hemoglobin are standard in the literature.

[49] We also estimated weight-for-height and an indicator of whether the child is wasted (weight-for-height lower than -2), and found similar results.

allocations are consistent with Pareto efficiency. To test these hypotheses I use the technology developed in the intra-household literature.

This work extends on previous work by showing how results differ when we increase the spatial dispersion among family members, including families with members spread across international borders. Presumably, families can coordinate more easily when everyone lives close to each other relative to families with members living far away from each other. With prevalent migration from Mexico to the United States, Mexico constitutes an ideal setting to look at these questions, and I use extremely rich population-representative longitudinal data that was designed to follow baseline respondents within Mexico and into the United States. Following the literature, I identify family members through household split-offs and take household formation and location decisions as given.

I test these hypotheses on two sets of outcomes: household budget shares and child human capital indicators. Household budget shares constitute a natural outcome of interest when analyzing allocation decisions. However, expenditures are arguably difficult to observe across households, especially when these households are spatially dispersed. Additionally, if coordination requires effort, families might not be willing to coordinate over all dimensions. With these ideas in mind, I also evaluate family behavior when it comes to investing in the future generations, i.e. investments in child human capital. In particular, I look at child height as a marker of nutrition, years of education, and raven scores as a marker of cognition. These outcomes are arguably very salient to family members and more easily observable. Household resources and resources from non-co-resident family members are measured by the log of per-capita expenditures.

Results suggest that the combination of looking at different levels of spatial dispersion and analyzing different dimensions of family behavior delivers very interesting results that would have not been evident had we pooled all families together or had we only analyzed one set of outcomes.

When I evaluate household budget shares I find that resources in the hands of non-co-resident family members do influence household budget allocations, but I reject the unitary model of the family. In other words, resources from within and outside the household have different effects on household allocation decisions. In terms of whether families are Pareto efficient, the conclusion depends on how we want to interpret the results. When I run the joint test that the ratio of marginal effects is the same across any pair of goods, I fail to reject it is. Therefore, this is evidence consistent with Pareto efficiency. However, if I count the number of individual tests that are significant at the 5% level, there are more cases than what we would expect to see at random. All three conclusions apply to all family types, regardless of the spatial dispersion among their members.

When I extend the analysis to child human capital indicators a number of differences emerge. For the two groups of families that have all households in Mexico, family resources from outside the household have a significant impact on child human capital. Furthermore, their effect is indistinguishable from the effect of within-household resources. In other words, the unitary model of the family is not rejected. Not surprisingly, I then fail to reject Pareto efficiency. In contrast, when I look at children in families with members spread across Mexico and the U.S., resources from non-co-resident family members do not have a significant effect on child human capital. This result is robust to a number of specifications, including: differentiating between resources originated in Mexico or in the U.S., using measures of resources that take into account inter-household transfers, implementing non-linear specifications, and extending the analysis to other health indicators.

In sum, this analysis provides with additional evidence on the importance of looking outside the household unit, and take into account interactions among non-co-resident family members. It contributes to the existing literature by highlighting important differences across families with various degrees of spatial dispersion among their members, as well as across different dimensions of family behavior. Two main findings emerge from this analysis. First, families seem to be able to achieve cooperation even when spread across international borders. Second, families with all households in

Mexico seem to behave as a single unit when it comes to investing in child human capital.

2.10 Primary tables

Table 15: MxFLS Structure

		HOUSEHOLDS		FAMILIES	
MxFLS1					
At baseline		8,440		8,440	
At least one member alive*					
	By MxFLS2	8,386		8,386	
	By MxFLS3	8,335		8,335	
MxFLS2					
All households in sample				#	%
	In MX	8,437			
	In US	503			
	Total	**8,940**	**Total**	**7,563**	**90%**
Households with extended family in sample					
	In MX	1,965	All hhs in Mx	692	
	In US	503	At least one hh in US	425	
	Total	**2,468**	**Total**	**1,117**	
MxFLS3					
All households in sample				#	%
	In MX	9,202			
	In US	740			
	Total	**9,942**	**Total**	**7,210**	**87%**
Households with extended family in sample					
	In MX	3,992	All hhs in Mx	1,424	
	In US	688	At least one hh in US	535	
	Total	**4,680**	**Total**	**1,959**	

* Eliminate households for which all members have died by the time of the follow-up

Table 16: Evidence of help provided across households within families

Family group	Sample of families								Test differences between[+] …			
	All families	**Same locality *neighbor***	**Spread across Mexico**			**Spread across Mx & US *international families***			**Neighbor Across Mx**	**Across Mx Int.**	**Across Mx: Orig/New**	**Int: Mx/US**
Location of interview			**All**	**in orig. loc**	**in new loc**	**All**	**in Mx**	**in US**				
Households												
Provided and/or received help												
% Households …												
neither provided nor received		**0.45**	**0.44**	0.41	0.47	**0.29**	0.35	0.22				
only received help		**0.12**	**0.11**	0.14	0.09	**0.13**	0.21	0.04				
only provided help		**0.18**	**0.19**	0.2	0.19	**0.35**	0.15	0.57				
both		**0.26**	**0.26**	0.26	0.26	**0.23**	0.29	0.17				
Provided help												
% Households sent												
$/gifts	0.48	**0.44**	**0.45**	0.46	0.44	**0.58**	0.44	0.74	0.46	0.00	0.60	0.00
time*	0.12	**0.12**	**0.11**	0.11	0.09	**0.11**	0.11	N/A	0.22	0.72	0.00	N/A
Amount sent												
Unconditional	1,976	**1,922**	**1,996**	1,837	2,161	**2,056**	1,767	2,370	0.71	0.77	0.35	0.00
Conditional on sending	4,146	**4,476**	**4,476**	4,058	4,921	**3,550**	4,075	3,214	1.00	0.02	0.22	0.02
Share over (cond. on sending)												
Total expenditures	37.91	**44.38**	**39.64**	36.86	42.71	**28.03**	44.09	16.17	0.26	0.00	0.40	0.00
Total income	48.48	**55.28**	**51.30**	48.68	54.18	**37.64**	68.28	18.91	0.41	0.00	0.49	0.00
Received help												
% Households received												
$/gifts	0.37	**0.37**	**0.37**	0.40	0.35	**0.36**	0.50	0.21	0.99	0.47	0.10	0.00
time*	0.08	**0.09**	**0.07**	0.07	0.08	**0.08**	0.08	N/A	0.18	0.88	0.00	0.00
Amount received												
Unconditional	1,789	**1,650**	**1,774**	1,722	1,827	**2,039**	3,823	128	0.52	0.26	0.75	0.00
Conditional on sending	4,915	**4,496**	**4,803**	4,385	5,294	**5,747**	7,735	622	0.51	0.10	0.27	0.00
Share over (cond. on receiving)												
Total expenditures	55.35	**47.09**	**49.31**	55.02	42.84	**75.54**	99.42	2.74	0.69	0.00	0.21	0.00
Total income	84.04	**77.90**	**78.60**	88.44	66.85	**100.48**	135.63	2.19	0.94	0.05	0.16	0.00
Net help: amount sent - amount received												
Unconditional	214	**292**	**221**	117	328	**74**	-1,955	2,255	0.78	0.61	0.64	0.00
Conditional on sending or receiving	361	**539**	**396**	199	620	**104**	-3,041	2,913	0.76	0.54	0.61	0.00

	Sample of families								Test differences between[+] …			
Family group	**All families**	**Same locality** ***neighbor***	**Spread across Mexico**			**Spread across Mx & US** ***international families***			**Neighbor Across Mx**	**Across Mx Int.**	**Across Mx: Orig/New**	**Int: Mx/US**
Location of interview			**All**	**in orig. loc**	**in new loc**	**All**	**in Mx**	**in US**				
Individuals												
Provided help to …												
children (cond on ch outside the hh)												
$/gifts	0.18	**0.18**	**0.18**	0.16	0.21	**0.18**	0.14	0.28	0.99	0.98	0.28	0.00
time*	0.02	**0.03**	**0.02**	0.02	0.01	**0.02**	0.02	N/A	…	…	…	N/A
parent (cond on parent outside the hh)												
$/gifts	0.31	**0.27**	**0.28**	0.26	0.29	**0.45**	0.26	0.79	0.52	0.00	0.28	0.00
time*	0.05	**0.06**	**0.05**	0.05	0.05	**0.05**	0.05	N/A	…	…	…	N/A
sibling (Mx: cond - US: uncond.)												
$/gifts	0.09	**0.08**	**0.08**	0.06	0.10	**0.12**	0.07	0.23	0.83	0.00	0.01	0.00
time*	0.02	**0.02**	**0.02**	0.02	0.02	**0.02**	0.02	N/A	…	…	…	N/A
other (uncond)												
$/gifts	0.06	**0.05**	**0.06**	0.05	0.07	**0.08**	0.05	0.19	0.07	0.01	0.11	0.00
time*	0.01	**0.01**	**0.01**	0.01	0.01	**0.00**	0.00	N/A	…	…	…	N/A
other (cond on having sent some transfers)												
$/gifts	0.20	**0.19**	**0.22**	0.22	0.21	**0.20**	0.21	0.19	0.15	0.30	0.79	0.50
time*	0.12	**0.12**	**0.14**	0.15	0.12	**0.09**	0.09	N/A	…	…	…	N/A
spouse (US sample)												
$/gifts	0.87	**N/A**	**N/A**	N/A	N/A	**0.87**	N/A	0.87	N/A	N/A	N/A	N/A

Notes: Shares conditional on having sent/received help. All magnitudes correspond to the 12 months previos to the interview date.

* Help in time only available for households interviewed in Mexico.

[+] p-values associated with the test that means are statistically different, taking into account clustering at the family level.

Source: MxFLS3

Table 17: Summary statistics of main variables used in the analysis, by family type

	All families			Same locality *neighbor families*			Spread across Mexico			Spread across Mx & US *international families*		
	mean	sd	median	mean	sd	median	mean	sd	median	mean	sd	median
Sample of households												
food share	50.05	15.95	50.43	51.41	15.32	51.76	50.70	16.39	51.10	46.00	16.53	45.53
personal goods share	11.50	8.90	9.56	11.44	8.75	9.55	11.87	9.41	9.94	11.36	8.86	9.38
semi-durables share	5.30	6.68	3.11	5.20	6.46	3.04	5.60	7.39	3.17	5.33	6.67	3.26
transport and communication sh	9.21	9.81	6.83	8.04	8.83	5.71	8.36	8.97	6.28	12.88	11.76	11.25
housing share	23.95	13.06	21.58	23.91	12.99	21.53	23.47	12.85	21.01	24.44	13.41	22.15
log household pce	5.45	0.87	5.41	5.31	0.75	5.31	5.39	0.83	5.38	5.81	1.03	5.84
log family pce	5.46	0.65	5.45	5.31	0.63	5.29	5.40	0.66	5.33	5.86	0.52	5.85
log extended-family pce	5.45	0.81	5.43	5.30	0.72	5.29	5.39	0.79	5.37	5.83	0.88	5.86
# observations		4,371			2,677			756			938	
Sample of children												
height-for-age												
height-for-age	-0.46	1.25	-0.48	-0.54	1.21	-0.54	-0.42	1.32	-0.41	-0.24	1.30	-0.29
log household pce	5.15	0.77	5.13	5.08	0.70	5.08	5.15	0.74	5.15	5.35	0.95	5.29
log family pce	5.26	0.61	5.25	5.15	0.59	5.13	5.19	0.60	5.18	5.68	0.47	5.71
log extended-family pce	5.33	0.75	5.29	5.21	0.71	5.21	5.20	0.66	5.19	5.85	0.75	5.82
interviewed in MxFLS2	0.21	0.41	0.00	0.25	0.43	0.00	0.12	0.33	0.00	0.15	0.36	0.00
interviewed in the U.S.	0.08	0.28	0.00	0.00	0.00	0.00	0.02	0.14	0.00	0.39	0.49	0.00
mother in household	0.99	0.12	1.00	0.98	0.13	1.00	0.99	0.09	1.00	0.99	0.10	1.00
father in household	0.85	0.35	1.00	0.86	0.35	1.00	0.87	0.33	1.00	0.82	0.39	1.00
mother's height	155	6	154	154	6	154	155	6	154	155	6	154
mother's education	8.52	3.27	9.00	8.52	3.35	9.00	8.84	3.22	9.00	8.25	3.02	9.00
household size	4.97	2.41	4.00	4.95	2.34	4.00	4.95	2.94	4.00	5.05	2.12	4.00
number of kids in household	2.18	1.22	2.00	2.16	1.22	2.00	2.18	1.28	2.00	2.28	1.18	2.00
family size	11.40	4.21	11.00	11.00	3.96	10.00	11.99	4.18	12.00	12.14	4.80	11.00
number of kids in family	4.11	2.18	4.00	3.99	2.18	4.00	4.47	2.22	4.00	4.18	2.10	4.00
rural	0.37	0.48	0.00	0.37	0.48	0.00	0.39	0.49	0.00	0.36	0.48	0.00
# observations		2,071			1,289			366			416	

	All families			Same locality *neighbor families*			Spread across Mexico			Spread across Mx & US *international families*		
	mean	sd	median	mean	sd	median	mean	sd	median	mean	sd	median
Sample of children												
years of education												
years of education	6.35	3.08	4.00	4.48	3.08	4.00	4.46	3.13	4.00	4.81	3.06	5.00
log household pce	5.18	0.77	4.95	4.93	0.70	4.94	4.94	0.78	5.02	5.01	0.90	4.98
log family pce	5.35	0.62	5.25	5.08	0.59	5.07	5.19	0.63	5.16	5.67	0.48	5.70
log extended-family pce	5.45	0.83	5.47	5.23	0.70	5.22	5.38	0.78	5.38	6.17	0.77	6.13
interviewed in MxFLS2	0.34	0.46	0.00	0.33	0.47	0.00	0.18	0.38	0.00	0.34	0.47	0.00
interviewed in the U.S.	0.10	0.23	0.00	0.00	0.00	0.00	0.02	0.14	0.00	0.20	0.40	0.00
mother in household	0.46	0.32	1.00	0.89	0.32	1.00	0.88	0.33	1.00	0.88	0.32	1.00
father in household	0.37	0.48	1.00	0.64	0.48	1.00	0.67	0.47	1.00	0.64	0.48	1.00
mother's height	153	6	153	153	6	153	153	6	153	152	6	153
mother's education	5.96	3.69	6.00	6.51	3.73	6.00	6.36	3.78	6.00	5.80	3.49	6.00
household size	4.98	2.57	5.00	5.86	2.62	5.00	5.65	2.69	5.00	5.92	2.37	5.00
number of kids in household	1.64	1.52	2.00	2.41	1.59	2.00	2.25	1.37	2.00	2.47	1.42	2.00
family size	10.36	4.25	10.00	11.08	4.21	10.00	11.10	4.25	10.00	11.02	4.35	10.00
number of kids in family	3.29	2.27	4.00	4.14	2.36	4.00	4.12	2.28	4.00	3.79	2.00	4.00
rural	0.38	0.50	0.00	0.39	0.49	0.00	0.45	0.50	0.00	0.52	0.50	1.00
# observations		2,699			2,176			614			909	

	All families			Same locality *neighbor families*			Spread across Mexico			Spread across Mx & US *international families*		
	mean	sd	median	mean	sd	median	mean	sd	median	mean	sd	median
Sample of children												
raven score												
raven score	51.17	21.30	55.56	56.40	21.56	58.33	53.96	20.06	55.56	54.47	21.34	55.56
log household pce	5.18	0.70	4.93	4.95	0.68	4.96	4.96	0.73	5.03	4.74	0.73	4.78
log family pce	5.35	0.59	5.23	5.10	0.58	5.09	5.18	0.54	5.18	5.62	0.48	5.63
log extended-family pce	5.45	0.83	5.47	5.24	0.70	5.23	5.35	0.70	5.40	6.32	0.69	6.25
interviewed in MxFLS2	0.34	0.46	0.00	0.33	0.47	0.00	0.18	0.39	0.00	0.36	0.48	0.00
interviewed in the U.S.	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
mother in household	0.46	0.31	1.00	0.90	0.31	1.00	0.90	0.30	1.00	0.89	0.31	1.00
father in household	0.37	0.47	1.00	0.67	0.47	1.00	0.71	0.46	1.00	0.65	0.48	1.00
mother's height	153	6	153	153	6	153	153	6	153	152	6	153
mother's education	5.96	3.68	6.00	6.64	3.72	6.00	6.52	3.70	6.00	5.47	3.44	6.00
household size	4.98	2.53	5.00	5.75	2.51	5.00	5.64	2.61	5.00	6.18	2.50	6.00
number of kids in household	1.64	1.48	2.00	2.40	1.53	2.00	2.28	1.31	2.00	2.59	1.46	2.00
family size	10.36	4.17	10.00	11.02	4.12	10.00	11.08	4.20	10.00	11.06	4.31	10.00
number of kids in family	3.29	2.22	4.00	4.11	2.30	4.00	4.07	2.19	4.00	3.82	2.02	4.00
rural	0.38	0.50	0.00	0.38	0.49	0.00	0.43	0.50	0.00	0.63	0.48	1.00
# observations		3,504			2,176			553			775	

Notes: Households with non-missing expenditure. Children with non-missing height-for-age, years of education, raven score.

Source: MxFLS2 & MxFLS3

Table 18: Estimated effect of family resources on household budget shares

Family type	**All families**				
Budget share	**food** (at home & meals out)	**personal care clothing health education recreation**	**semi durables cleaning insurance repair**	**transport communication**	**housing** (rent & utilities)
log pce					
***household* (β1)**	-4.13	2.55	2.09	4.96	-5.48
	[0.538]***	[0.269]***	[0.224]***	[0.330]***	[0.419]***
***other family members* (β2)**	-1.30	-0.05	0.32	-0.20	1.22
	[0.393]***	[0.233]	[0.181]*	[0.237]	[0.330]***
Family Resources Matter					
Joint p-value (β2)			**0.00**		
Unitary Test					
β1 - β2	-2.83	2.60	1.77	5.16	-6.70
p-value	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
Pareto Efficiency Test					
β1 / β2	3.18	-56.01	6.47	-24.92	-4.49
Joint p-value⁺			**0.00**		
# observations	4,371	4,371	4,371	4,371	4,371

Notes: Additional controls include household size and composition (number of children, prime-age males, senior males, senior females, with prime-age females as the omitted category); age, gender and years of education of household head; household location (dummy of whether the household is in the US, region within the U.S., state within Mexico and whether the place is rural if the household is in Mexico); interview date (quarter-year dummies); MxFLS wave (dummy if observation from MxFLS2, and interaction of U.S. and MxFLS2); and family characteristics (age of family head; maximum years of education, height and cognitive score among adult males; maximum years of education, height and cognitive score among adult women; family size, number of children, prime-age males, senior females and senior males; log of per-capita wealth at baseline in 2002, its square, and its cube).

[+] Corresponds to last row in Table 32

Source: MxFLS2 & MxFLS3

Table 19: Estimated effect of family resources on household budget shares, by family-type - Main specification

Family type	**All families**					**Same locality** *neighbor families*				
Budget share	**food** (at home & meals out)	**p. care clothing health education recreation**	**semi dur. cleaning insurance repair**	**transport comm.**	**housing** (rent & utilities)	**food** (at home & meals out)	**p. care clothing health education recreation**	**semi dur. cleaning insurance repair**	**transport comm.**	**housing** (rent & utilities)
log pce										
household **(β1)**										
spline - 0025	-2.42	2.92	1.69	4.83	-7.01	-2.89	3.72	1.45	4.39	-6.66
	[0.935]***	[0.472]***	[0.294]***	[0.414]***	[0.756]***	[1.302]**	[0.583]***	[0.383]***	[0.475]***	[1.066]***
spline - 2575	-8.94	1.98	2.78	7.22	-3.03	-7.99	1.67	2.57	6.54	-2.79
	[0.943]***	[0.527]***	[0.448]***	[0.596]***	[0.801]***	[1.149]***	[0.698]**	[0.578]***	[0.698]***	[1.042]***
spline - 7500	0.08	2.88	1.71	2.16	-6.82	-3.82	3.99	2.33	3.68	-6.18
	[1.322]	[0.665]***	[0.652]***	[0.835]***	[0.817]***	[1.795]**	[1.274]***	[0.795]***	[1.223]***	[1.262]***
other family members **(β2)**										
spline - 0025	-1.53	-0.30	0.22	-0.45	2.06	-1.58	-0.33	-0.23	0.06	2.08
	[0.875]*	[0.534]	[0.334]	[0.509]	[0.758]***	[1.110]	[0.701]	[0.412]	[0.606]	[1.002]**
spline - 2575	-1.12	-0.22	0.19	-0.12	1.26	-0.93	-0.64	0.43	0.56	0.58
	[0.856]	[0.500]	[0.414]	[0.531]	[0.754]*	[1.105]	[0.644]	[0.514]	[0.670]	[0.969]
spline - 7500	-1.11	0.50	0.62	-0.13	0.11	-2.22	2.03	0.11	-1.10	1.18
	[0.872]	[0.506]	[0.462]	[0.583]	[0.741]	[1.626]	[1.033]**	[0.669]	[0.918]	[1.333]

Family type	All families					Same locality *neighbor families*				
Budget share	**food** (at home & meals out)	**p. care clothing health education recreation**	**semi dur. cleaning insurance repair**	**transport comm.**	**housing** (rent & utilities)	**food** (at home & meals out)	**p. care clothing health education recreation**	**semi dur. cleaning insurance repair**	**transport comm.**	**housing** (rent & utilities)
Non-linearity Test										
p-value β1	0.00	0.49	0.21	0.00	0.00	0.03	0.13	0.21	0.05	0.04
p-value β2	0.92	0.44	0.74	0.88	0.16	0.85	0.10	0.65	0.41	0.62
Family Resources Matter										
p-value β2	0.06	0.28	0.22	0.61	0.00	0.22	0.08	0.43	0.74	0.01
Joint p-value			**0.00**					**0.04**		
Unitary Test										
β1 - β2										
spline - 0025	-0.89	3.21	1.46	5.28	-9.07	-1.31	4.04	1.68	4.34	-8.75
spline - 2575	-7.82	2.19	2.59	7.33	-4.29	-7.07	2.31	2.15	5.97	-3.37
spline - 7500	1.18	2.39	1.09	2.28	-6.94	-1.60	1.95	2.22	4.78	-7.36
p-value										
spline - 0025	0.54	0.00	0.00	0.00	0.00	0.50	0.00	0.01	0.00	0.00
spline - 2575	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.01	0.00	0.03
spline - 7500	0.49	0.01	0.24	0.03	0.00	0.56	0.28	0.05	0.01	0.00
Joint p-value	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
Pareto Efficiency Test										
β1 / β2										
spline - 0025	1.58	-9.82	7.56	-10.69	-3.41	1.83	-11.33	-6.33	78.83	-3.20
spline - 2575	7.97	-9.19	14.68	-61.15	-2.39	8.62	-2.61	6.02	11.61	-4.81
spline - 7500	-0.07	5.80	2.76	-17.26	-59.84	1.72	1.96	20.65	-3.33	-5.24
Joint p-value^{+}			**0.13**					**0.44**		
# observations	4,371	4,371	4,371	4,371	4,371	2,677	2,677	2,677	2,677	2,677

Family type	Spread across Mexico					Spread across Mx & US *international families*				
Budget share	**food** (at home & meals out)	**p. care clothing health education recreation**	**semi dur. cleaning insurance repair**	**transport comm.**	**housing** (rent & utilities)	**food** (at home & meals out)	**p. care clothing health education recreation**	**semi dur. cleaning insurance repair**	**transport comm.**	**housing** (rent & utilities)
log pce										
***household* (β1)**										
spline - 0025	-1.88	2.81	1.24	4.92	-7.09	-1.60	0.87	2.44	5.24	-6.95
	[2.186]	[1.444]*	[0.730]*	[1.265]***	[1.641]***	[1.916]	[1.001]	[0.727]***	[1.151]***	[1.521]***
spline - 2575	-8.45	0.78	3.06	9.07	-4.47	-11.28	4.26	2.12	8.13	-3.22
	[2.345]***	[1.380]	[1.081]***	[1.374]***	[2.062]**	[2.420]***	[1.457]***	[1.014]**	[1.787]***	[1.593]**
spline - 7500	2.13	1.14	-0.45	2.51	-5.32	2.64	2.17	2.17	1.23	-8.21
	[3.518]	[1.751]	[1.857]	[1.553]	[1.985]***	[2.205]	[0.906]**	[1.056]**	[1.508]	[1.346]***
***other family members* (β2)**										
spline - 0025	-3.71	0.12	2.13	-0.17	1.63	1.61	-1.23	1.06	-3.78	2.36
	[2.203]*	[1.052]	[0.785]***	[1.214]	[1.368]	[2.030]	[1.214]	[1.127]	[1.936]*	[2.087]
spline - 2575	1.15	-0.17	-0.89	-0.55	0.45	-4.11	0.28	-0.47	0.25	4.05
	[2.262]	[1.275]	[1.205]	[1.291]	[1.832]	[2.138]*	[1.267]	[1.040]	[1.609]	[1.908]**
spline - 7500	-1.92	1.46	0.93	1.17	-1.63	-0.06	-1.47	0.40	0.67	0.46
	[2.594]	[1.504]	[2.169]	[1.220]	[1.636]	[1.236]	[0.759]*	[0.578]	[0.956]	[1.097]

p-value β1	Spread across Mexico					Spread across Mx & US *international families*				
Family Resources Matter	**food**	**p. care**	**semi dur.**	**transport**	**housing**	**food**	**p. care**	**semi dur.**	**transport**	**housing**
Non-linearity Test										
p-value β1	0.11	0.58	0.39	0.02	0.59	0.00	0.25	0.96	0.01	0.06
p-value β2	0.40	0.72	0.13	0.60	0.30	0.26	0.58	0.72	0.10	0.23
Family Resources Matter										
p-value β2	0.07	0.39	0.12	0.14	0.37	0.14	0.20	0.29	0.17	0.01
Joint p-value			**0.11**					**0.01**		
Unitary Test										
β1 - β2										
spline - 0025	1.83	2.69	-0.90	5.09	-8.72	-3.21	2.10	1.38	9.03	-9.30
spline - 2575	-9.59	0.94	3.95	9.62	-4.92	-7.17	3.98	2.59	7.88	-7.27
spline - 7500	4.05	-0.32	-1.38	1.34	-3.69	2.70	3.64	1.78	0.57	-8.67
p-value										
spline - 0025	0.61	0.13	0.43	0.01	0.00	0.25	0.20	0.33	0.00	0.00
spline - 2575	0.01	0.65	0.03	0.00	0.10	0.03	0.04	0.07	0.00	0.00
spline - 7500	0.43	0.89	0.71	0.53	0.19	0.30	0.00	0.16	0.75	0.00
Joint p-value	**0.04**	**0.31**	**0.08**	**0.00**	**0.00**	**0.02**	**0.00**	**0.02**	**0.00**	**0.00**
Pareto Efficiency Test										
β1 / β2										
spline - 0025	0.51	24.22	0.58	-29.84	-4.35	-1.00	-0.71	2.31	-1.39	-2.95
spline - 2575	-7.36	-4.67	-3.44	-16.65	-9.83	2.74	15.31	-4.56	32.91	-0.79
spline - 7500	-1.11	0.78	-0.49	2.15	3.26	-47.83	-1.47	5.49	1.85	-17.81
Joint p-value+			**0.78**					**0.32**		
# observations	756	756	756	756	756	938	938	938	938	938

Notes: Additional controls include household size and composition (number of children, prime-age males, senior males, senior females, with prime-age females as the omitted category); age, gender and years of education of household head; household location (dummy of whether the household is in the US, region within the U.S., state within Mexico and whether the place is rural if the household is in Mexico); interview date (quarter-year dummies); MxFLS wave (dummy if observation from MxFLS2, and interaction of U.S. and MxFLS2); and family characteristics (age of family head; maximum years of education, height and cognitive score among adult males; maximum years of education, height and cognitive score among adult women; family size, number of children, prime-age males, senior females and senior males; log of per-capita wealth at baseline in 2002, its square, and its cube).

+ Corresponds to last row in Table 20

Source: MxFLS2 & MxFLS3

Table 20: Pareto efficiency test when analyzing household budget shares - Main specification

Family type		All families			Same locality			Spread across Mx			Spread across Mx & US		
Individual tests of pairwise ratios													
household in …		**spl1**	**spl1**	**spl1**	**spl1**	**spl1**	**spl1**	**spl1**	**spl1**	**spl1**	**spl1**	**spl1**	**spl1**
extended family in …		**spl1**	**spl2**	**spl3**	**spl1**	**spl2**	**spl3**	**spl1**	**spl2**	**spl3**	**spl1**	**spl2**	**spl3**
food	personal goods	0.03	0.21	0.59	0.07	0.27	0.61	0.12	0.25	0.24	0.85	0.12	0.88
	semi-durables	0.20	0.56	0.79	0.11	0.68	0.33	0.91	0.55	0.70	0.27	0.14	0.17
	transport/comm.	0.03	0.40	0.40	0.15	0.74	0.10	0.04	0.10	0.24	0.83	0.05	0.17
	housing	0.03	0.32	0.57	0.07	0.47	0.25	0.07	0.13	0.23	0.64	0.04	0.19
personal goods	semi-durables	0.37	0.24	0.21	0.83	0.31	0.50	0.10	0.67	0.64	0.23	0.07	0.06
	transport/comm.	0.97	0.25	0.51	0.67	0.11	0.03	0.84	0.22	0.89	0.66	0.95	0.08
	housing	0.29	0.69	0.73	0.25	0.51	0.03	0.46	0.14	0.70	0.41	0.31	0.02
semi-durables	transport/comm.	0.31	0.67	0.09	0.60	0.89	0.29	0.01	0.26	0.64	0.04	0.79	0.33
	housing	0.05	0.28	0.12	0.61	0.48	0.46	0.01	0.21	0.52	0.11	0.26	0.37
transport/comm.	housing	0.12	0.21	0.57	0.07	0.15	0.56	0.45	0.80	0.37	0.28	0.12	0.73
household in …		**spl2**	**spl2**	**spl2**	**spl2**	**spl2**	**spl2**	**spl2**	**spl2**	**spl2**	**spl2**	**spl2**	**spl2**
extended family in …		**spl1**	**spl2**	**spl3**	**spl1**	**spl2**	**spl3**	**spl1**	**spl2**	**spl3**	**spl1**	**spl2**	**spl3**
food	personal goods	0.41	0.33	0.63	0.46	0.17	0.18	0.85	0.96	0.52	0.92	0.35	0.39
	semi-durables	0.12	0.71	0.15	0.03	0.81	0.68	0.81	0.68	0.87	0.44	0.21	0.09
	transport/comm.	0.01	0.16	0.43	0.26	0.84	0.05	0.02	0.76	0.72	0.08	0.12	0.12
	housing	0.01	0.08	0.75	0.06	0.46	0.43	0.09	0.95	0.85	0.28	0.02	0.18
personal goods	semi-durables	0.81	0.54	0.57	0.44	0.21	0.17	0.92	0.96	0.67	0.53	0.60	0.07
	transport/comm.	0.60	0.73	0.43	0.96	0.25	0.03	0.54	0.92	0.49	0.21	0.92	0.08
	housing	0.03	0.34	0.98	0.17	0.70	0.04	0.41	0.94	0.65	0.21	0.06	0.03
semi-durables	transport/comm.	0.68	0.62	0.10	0.23	0.74	0.31	0.06	0.58	0.78	0.18	0.63	0.30
	housing	0.02	0.08	0.53	0.46	0.33	0.50	0.09	0.71	0.97	0.19	0.24	0.54
transport/comm.	housing	0.01	0.10	0.41	0.05	0.39	0.99	0.22	0.91	0.64	0.89	0.03	0.46

Family type			All families			Same locality			Spread across Mx			Spread across Mx & US	
Individual tests of pairwise ratios													
household in …		**spl3**	**spl3**	**spl3**	**spl3**	**spl3**	**spl3**	**spl3**	**spl3**	**spl3**	**spl3**	**spl3**	**spl3**
extended family in …		**spl1**	**spl2**	**spl3**	**spl1**	**spl2**	**spl3**	**spl1**	**spl2**	**spl3**	**spl1**	**spl2**	**spl3**
food	personal goods	0.06	0.25	0.21	0.28	0.08	0.88	0.89	0.93	0.39	0.65	0.09	0.41
	semi-durables	0.09	0.53	0.26	0.04	0.79	0.25	0.29	0.29	0.82	0.91	0.10	0.74
	transport/comm.	0.07	0.38	0.25	0.36	0.93	0.08	0.32	0.40	0.39	0.29	0.46	0.61
	housing	0.04	0.31	0.20	0.08	0.24	0.17	0.12	0.17	0.38	0.43	0.10	0.93
personal goods	semi-durables	0.82	0.21	0.58	0.29	0.28	0.28	0.71	0.42	0.67	0.55	0.87	0.09
	transport/comm.	0.33	0.63	0.50	0.96	0.20	0.06	0.58	0.40	0.59	0.07	0.87	0.32
	housing	0.04	0.98	0.29	0.11	0.70	0.02	0.53	0.15	0.47	0.48	0.24	0.07
semi-durables	transport/comm.	0.42	0.63	0.30	0.39	0.80	0.38	0.64	0.96	0.65	0.08	0.92	0.68
	housing	0.12	0.16	0.17	0.42	0.23	0.49	0.13	0.55	0.67	0.17	0.20	0.38
transport/comm.	housing	0.73	0.56	0.88	0.14	0.15	0.72	0.61	0.67	0.78	0.06	0.60	0.42
Simultaneous tests of pairwise ratios													
household in …	**extended family in …**												
spl1 (0-25)	spl1		**0.29**			**0.54**			**0.09**			**0.44**	
	spl2		**0.69**			**0.60**			**0.65**			**0.51**	
	spl3		**0.67**			**0.27**			**0.86**			**0.18**	
spl2 (25-75)	spl1		**0.13**			**0.32**			**0.37**			**0.61**	
	spl2		**0.63**			**0.83**			**1.00**			**0.31**	
	spl3		**0.65**			**0.25**			**0.95**			**0.20**	
spl3 (75-00)	spl1		**0.38**			**0.45**			**0.67**			**0.47**	
	spl2		**0.78**			**0.60**			**0.77**			**0.67**	
	spl3		**0.72**			**0.26**			**0.95**			**0.48**	
All ratios			**0.13**			**0.44**			**0.78**			**0.32**	

Notes: p-values associated with PE test based on model in Table 19

Table 21: Estimated effect of family resources on child human capital indicators, by family-type – Main specifiacion

Family type	**All families**			**Same locality** *neighbor families*			**Spread across Mexico**			**Spread across Mx & US** *international families*		
Child outcomes	**height-for-age**	**years of education**	**cognitive score (%)**	**height-for-age**	**years of education**	**cognitive score (%)**	**height-for-age**	**years of education**	**cognitive score (%)**	**height-for-age**	**years of education**	**cognitive score (%)**
log pce												
household **(β1)**	0.25	0.19	3.12	0.17	0.14	3.36	0.27	0.38	1.91	0.39	0.25	2.31
	[0.0527]***	[0.0482]***	[0.702]***	[0.0631]***	[0.0625]**	[0.956]***	[0.139]*	[0.0996]***	[1.429]	[0.126]***	[0.100]**	[1.263]*
other family members **(β2)**	0.15	0.09	1.71	0.18	0.14	1.92	0.27	0.19	3.21	0.03	-0.02	0.99
	[0.0513]***	[0.0419]**	[0.628]***	[0.0643]***	[0.0541]***	[0.838]**	[0.146]*	[0.0939]**	[1.416]**	[0.0995]	[0.0750]	[1.451]
Family Resources Matter												
Joint p-value (β2)		**0.00**			**0.00**			**0.00**			**0.81**	
Unitary Test												
β1 - β2	0.10	0.10	1.41	-0.01	-0.01	1.44	0.01	0.18	-1.30	0.37	0.27	1.32
p-value	0.21	0.16	0.18	0.90	0.95	0.32	0.98	0.22	0.58	0.05	0.05	0.51
Pareto Efficiency Test												
β1 / β2	1.70	2.05	1.82	0.93	0.96	1.75	1.02	1.96	0.60	15.14	-11.47	2.33
Joint p-value⁺		**0.99**			**0.75**			**0.60**			**0.86**	
# observations	2,071	3,699	3,504	1,289	2,176	2,176	366	614	553	416	909	775

Notes: Additional controls include age, sex, household size and number of children under 15, family size, number of children under 15 and number of female adulst in family, location (U.S. and Mexican state dummies), MxFLS2 dummy, MxFLS2 interacted with U.S. dummy, interview date.

[+] Corresponds to last row in Table 22

Source: MxFLS2 & MxFLS3

Table 22: Pareto efficiency test when analyzing child human capital indicators - Main specification

Family type		**All families**	**Same locality**	**Spread across Mx**	**Spread across Mx & US**
Individual tests of pairwise ratios					
height-for-age	yrs. of education	0.76	0.93	0.50	0.67
	cognitive score	0.82	0.34	0.66	0.61
yrs. of education	cognitive score	0.93	0.36	0.20	0.43
Simultaneous tests of pairwise ratios					
All Ratios		**0.99**	**0.75**	**0.60**	**0.86**

Notes: p-values associated with PE test based on model in Table 21

Table 23: Estimated effect of family resources on child human capital indicators among international families, by location of the child and location of resources

Family type	All families			Spread across Mx & US *international families*		
Child outcomes	height-for-age	years of education	cognitive score (%)	height-for-age	years of education	cognitive score (%)
PANEL A: interactions w/location of the child						
log pce						
***household* (β1)**						
child in Mexico	0.23	0.19		0.39	0.25	
	[0.0540]***	[0.0485]***		[0.144]***	[0.105]**	
child in U.S.	0.44	0.24		0.38	0.13	
	[0.173]**	[0.212]		[0.192]**	[0.204]	
***other family members* (β2)**						
child in Mexico	0.16	0.10		-0.09	-0.04	
	[0.0551]***	[0.0431]**		[0.164]	[0.0858]	
child in U.S.	0.07	-0.02		0.12	0.05	
	[0.108]	[0.123]		[0.116]	[0.145]	
Test equal marginal effects across locations						
p-value (β1)	0.26	0.80		0.96	0.62	
p-value (β2)	0.46	0.34		0.29	0.60	
PANEL B: location of resources						
log pce						
***household* (β1)**						
resources in Mexico	0.23	0.18	3.10	0.41	0.25	2.42
	[0.0537]***	[0.0481]***	[0.711]***	[0.148]***	[0.106]**	[1.300]*
resources in the U.S.	0.40	0.19	0.00	0.34	0.12	0.00
	[0.173]**	[0.219]	0.00	[0.195]*	[0.213]	0.00
***other family members* (β2)**						
resources in Mexico	0.183	0.138	1.630	0.215	0.089	-0.245
	[0.0532]***	[0.0448]***	[0.684]**	[0.123]*	[0.118]	[1.847]
resources in the U.S.	-0.145	-0.140	0.220	-0.185	-0.113	0.320
	[0.152]	[0.0695]**	[1.177]	[0.151]	[0.0724]	[1.320]
Test equal marginal effects across resource sources						
p-value (β1)	0.33	0.95		0.74	0.60	
p-value (β2)	0.04	0.00	0.29	0.05	0.12	0.80
# observations	2,071	3,699	3,504	416	909	775

Notes: Additional controls include age, sex, household size and number of children under 15, family size, number of children under 15 and number of female adulst in family, location (U.S. and Mexican state dummies), MxFLS2 dummy, MxFLS2 interacted with U.S. dummy, interview date.
Source: MxFLS2 & MxFLS3

Table 24: Estimated effect of family resources on child human capital indicators among international families - Alternative measures of resources

Measure of resources	**Per-capita expenditure**+			**Expenditure net of transfers**++			**Expenditures plus transfers sent**+++		
Estimation	(IV)			(OLS)			(OLS)		
Child outcomes	**height-for-age**	**years of education**	**cognitive score (%)**	**height-for-age**	**years of education**	**cognitive score (%)**	**height-for-age**	**years of education**	**cognitive score (%)**
log pce									
household **(β1)**	0.233	0.374	10.97	0.04	0.21	3.01	0.09	0.14	2.71
	[0.330]	[0.231]	[4.354]**	[0.0940]	[0.0672]***	[1.155]***	[0.0880]	[0.0542]**	[1.045]**
other family members **(β2)**	0.375	-0.044	-5.694	-0.06	-0.05	-0.60	0.03	-0.06	-0.24
	[0.171]**	[0.157]	[4.141]	[0.0689]	[0.0636]	[1.105]	[0.0692]	[0.0921]	[1.388]
# observations	416	909	775	352	635	493	352	635	493

Notes: All models have controls specified in notes Table 21.

+ MxFLS2 & MxFLS3. Instruments: log per-capita income, log per-capita wealth, log per-capita wealth at baseline. The two F-statistics below 10 are for household resources in hfa, and for other family resources in cognitive scores (8.61 and 5.62 respectively)

++ Only MxFLS3. Resources: Expenditure plus transfers sent minus transfers received.

+++ Only MxFLS3.

Source: MxFlS2 & MxFLS3

2.11 Supplementary tables

Table 25: Household structure in Mexico and in the United States, by MxFLS round

Households in	United States						Mexico					
MxFLS wave	MxFLS2			MxFLS3			MxFLS2			MxFLS3		
Living arrangements												
	mean	**sd**	**median**	**mean**	**sd**	**median**	**mean**	**sd**	**median**	**mean**	**sd**	**median**
# individuals in dwelling	5.94	2.86	5	4.93	2.72	5	NA	NA	NA	NA	NA	NA
# relatives in dwelling	NA	NA	NA	3.77	2.44	4	NA	NA	NA	NA	NA	NA
household size	2.61	1.78	2	2.71	1.64	3	3.95	2.03	3	4.1	2.13	4
Total households		503			602			1,965			3,992	
Relationship to head of household												
	Freq.	**Percent**		**Freq.**	**Percent**		**Freq.**	**Percent**		**Freq.**	**Percent**	
Head	503	40.6		602	36.87		1,965	24.16		3,992	24.52	
Spouse	144	11.62		318	19.47		1,420	18.18		2,974	18.41	
Son/Daughter	183	14.77		558	34.17		3,123	39.99		6,363	39.38	
Step child	8	0.65		36	2.2		91	1.17		137	0.85	
Son/Daugher in law	10	0.81		9	0.55		189	2.42		428	2.65	
Parent	16	1.29		8	0.49		66	0.85		136	0.84	
Parent in law	6	0.48		7	0.43		24	0.31		101	0.63	
Sibling	94	7.59		31	1.9		54	0.69		130	0.8	
Sibling in law	32	2.58		14	0.86		37	0.47		82	0.51	
Grandchild	15	1.21		15	0.92		644	8.25		1,496	9.26	
Grandparent	2	0.16		0	0		3	0.04		18	0.11	
Uncle/Aunt	25	2.02		1	0.06		6	0.08		19	0.12	
Nephew/Niece	45	3.63		19	1.16		85	1.09		162	1	
Cousin	40	3.23		6	0.37		5	0.06		14	0.09	
Not relative	.	.		5	0.31		10	0.13		25	0.34	
Other	116	9.36		4	0.24		17	0.21		75	0.47	
Missing	0	0		0	0		71	1.92		6	0.04	
Total individuals	1,239	100		1,633	100		7,810	100		16,158	100	

Notes: Mexican sample restricted to families that split by second or third wave respectively. Sample sizes are smaller than in table 1 due to missing information on household composition.
Source: MxFLS2 & MxFLS3

Table 26: Family structure, by family type

Family group / Location of interview	Sample of families							Test differences between[+] …			
Family group	Same locality *Neighbor fam.*	Spread across Mexico			Spread across Mexico & U.S. *International families*			Neighbor vs. Across Mx	Across Mx vs. International	Across Mx: Orig vs. New	International: Mx vs. US
Location of interview		All	in orig loc	in new loc	All	in Mx	in US				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Household socio-demographic characteristics											
household size	**3.97**	**3.93**	4.36	3.47	**3.64**	4.21	2.54	**0.61**	**0.00**	0.00	0.00
# children 0-14	**1.21**	**1.21**	1.15	1.26	**1.04**	1.17	0.78	**0.93**	**0.00**	0.17	0.00
# adults (15+)	**2.76**	**2.72**	3.20	2.21	**2.49**	2.88	1.75	**0.50**	**0.00**	0.00	0.00
# prime age female adults	**1.25**	**1.23**	1.38	1.08	**1.09**	1.32	0.66	**0.71**	**0.00**	0.00	0.00
# prime age male adults	**1.15**	**1.13**	1.25	1.00	**1.09**	1.11	1.05	**0.53**	**0.27**	0.00	0.19
# senior females	**0.18**	**0.16**	0.25	0.07	**0.15**	0.22	0.01	**0.26**	**0.44**	0.00	0.00
# senior males	**0.18**	**0.19**	0.32	0.06	**0.16**	0.23	0.03	**0.38**	**0.05**	0.00	0.00
rural locality	**0.35**	**0.38**	0.51	0.25	**0.33**	0.51	N/A	**0.11**	**0.04**	0.00	N/A
household head […]											
age	**40.95**	**40.01**	48.22	31.25	**39.27**	43.71	30.74	**0.05**	**0.17**	0.00	0.00
female	**0.28**	**0.29**	0.22	0.36	**0.29**	0.39	0.10	**0.52**	**0.93**	0.00	0.00
years of education	**7.13**	**7.42**	6.01	8.92	**6.46**	5.64	8.02	**0.11**	**0.00**	0.00	0.00
married	**0.13**	**0.18**	0.17	0.20	**0.29**	0.12	0.63	**0.00**	**0.00**	0.30	0.00
Family socio-demographic characteristics											
# households	**2.23**	**2.34**	N/A	N/A	**2.50**	N/A	N/A	**0.00**	**0.00**	N/A	N/A
# individuals	**8.93**	**9.26**	N/A	N/A	**9.04**	N/A	N/A	**0.13**	**0.42**	N/A	N/A
# children 0-14	**2.71**	**2.84**	N/A	N/A	**2.57**	N/A	N/A	**0.29**	**0.05**	N/A	N/A
# adults (15+)	**6.14**	**6.36**	N/A	N/A	**6.19**	N/A	N/A	**0.11**	**0.33**	N/A	N/A
max. […] among male adults											
years of education	**9.43**	**9.26**	N/A	N/A	**8.87**	N/A	N/A	**0.11**	**0.00**	N/A	N/A
height	**168.4**	**168.1**	N/A	N/A	**167.6**	N/A	N/A	**0.12**	**0.03**	N/A	N/A
raven score	**65.96**	**65.91**	N/A	N/A	**64.94**	N/A	N/A	**0.93**	**0.16**	N/A	N/A
max. […] among female adults											
years of education	**9.56**	**9.35**	N/A	N/A	**8.94**	N/A	N/A	**0.03**	**0.00**	N/A	N/A
height	**157.0**	**156.8**	N/A	N/A	**155.9**	N/A	N/A	**0.37**	**0.00**	N/A	N/A
raven score	**66.23**	**66.36**	N/A	N/A	**64.27**	N/A	N/A	**0.81**	**0.00**	N/A	N/A

Family group	Same locality *Neighbor fam.*	Spread across Mexico			Spread across Mexico & U.S. *International families*			Neighbor vs. Across Mx	Across Mx vs. International	Across Mx: Orig vs. New	International: Mx vs. US
	Sample of families							**Test differences between+ …**			
Location of interview		**All**	**in orig loc**	**in new loc**	**All**	**in Mx**	**in US**				
	(1)	**(2)**	(3)	(4)	**(5)**	(6)	(7)	**(8)**	**(9)**	(10)	(11)
Family socio-demographic characteristics (continued)											
% of families with […]											
children 0-14	**0.91**	**0.89**	N/A	N/A	**0.81**	N/A	N/A	**0.30**	**0.00**	N/A	N/A
ch in hfa sample	**0.55**	**0.50**	N/A	N/A	**0.40**			**0.03**	**0.00**		
ch in edu sample	**0.68**	**0.70**	N/A	N/A	**0.71**	N/A	N/A	**0.40**	**0.62**	N/A	N/A
ch in cog sample	**0.68**	**0.65**	N/A	N/A	**0.63**	N/A	N/A	**0.30**	**0.55**	N/A	N/A
% households within family with […]											
children 0-14	**0.67**	**0.63**	N/A	N/A	**0.48**	N/A	N/A	**0.03**	**0.00**	N/A	N/A
ch in hfa sample	**0.31**	**0.26**	N/A	N/A	**0.18**	N/A	N/A	**0.00**	**0.00**	N/A	N/A
ch in edu sample	**0.40**	**0.38**	N/A	N/A	**0.36**	N/A	N/A	**0.35**	**0.13**	N/A	N/A
ch in cog sample	**0.40**	**0.36**	N/A	N/A	**0.29**	N/A	N/A	**0.02**	**0.00**	N/A	N/A
Number of […]											
ch in hfa sample	**0.85**	**0.78**	0.33	0.45	**0.61**	0.42	0.18	**0.19**	**0.00**		
ch in edu sample	**1.46**	**1.39**	0.92	0.47	**1.58**	1.33	0.25	**0.33**	**0.01**		
ch in cog sample	**1.44**	**1.27**	0.83	0.43	**1.34**	1.34	0.00	**0.01**	**0.33**		

Notes: + p-values associated with the test that means are statistically different, taking into account clustering at the family level.

Source: MxFLS2 & MxFLS3

Table 27: Difference in mean family resources and outcomes of the analysis, by family type

	Sample of families							Test differences between[+] …			
Family group	**Same locality** *Neighbor fam.*	**Spread across Mexico**			**Spread across Mexico & U.S.** *International families*			**Neighbor vs. Across Mx**	**Across Mx vs. International**	**Across Mx: Orig vs. New**	**International: Mx vs. US**
Location of interview		**All**	**in orig loc**	**in new loc**	**All**	**in Mx**	**in US**				
	(1)	**(2)**	(3)	(4)	**(5)**	(6)	(7)	**(8)**	**(9)**	(10)	(11)
Household resources											
log wealth in 2002	**11.13**	**10.98**	10.96	11.00	**10.90**	10.87	10.95	**0.20**	**0.53**	0.50	0.28
log total expenditure	**6.57**	**6.64**	6.53	6.77	**6.91**	6.63	7.44	**0.03**	**0.00**	0.00	0.00
log total income	**6.15**	**6.10**	6.11	6.10	**6.45**	6.03	7.21	**0.39**	**0.00**	0.87	0.00
log wealth	**8.89**	**8.82**	8.98	8.66	**8.92**	9.12	8.69	**0.49**	**0.35**	0.02	0.00
log per-capita expenditures	**5.31**	**5.39**	5.18	5.61	**5.81**	5.33	6.71	**0.04**	**0.00**	0.00	0.00
log per-capita income	**4.87**	**4.83**	4.73	4.94	**5.33**	4.70	6.48	**0.49**	**0.00**	0.01	0.00
log per-capita wealth	**7.59**	**7.52**	7.61	7.43	**7.61**	7.43	7.95	**0.47**	**0.38**	0.20	0.00
Correlation between household (pce) and other family resources (epe)[++]											
% hhs w/ pce & epe in same portion of the distribution	**0.48**	**0.44**	0.44	0.44	**0.25**	0.23	0.25				
correlation btw pce & epe	**0.43**	**0.32**	0.38	0.40	**-0.36**	0.07	0.03				
Outcomes analyzed in the paper											
Household budget shares											
food	**51.41**	**50.70**	52.73	48.53	**46.00**	52.81	35.48	**0.32**	**0.00**	0.00	0.00
personal goods	**11.44**	**11.87**	10.99	12.82	**11.36**	12.00	10.35	**0.23**	**0.20**	0.00	0.00
semi-durables	**5.20**	**5.60**	4.83	6.42	**5.33**	4.83	6.09	**0.15**	**0.37**	0.00	0.00
transport and comm.	**8.04**	**8.36**	7.95	8.80	**12.88**	9.35	18.33	**0.37**	**0.00**	0.13	0.00
housing	**23.91**	**23.47**	23.50	23.43	**24.44**	21.01	29.74	**0.41**	**0.10**	0.90	0.00
Child outcomes											
height-for-age	**-0.54**	**-0.42**	-0.54	-0.34	**-0.24**	-0.33	-0.09	**0.21**	**0.08**	0.19	0.07
years of education	**4.48**	**4.46**	4.99	3.48	**4.81**	4.99	4.07	**0.87**	**0.04**	0.00	0.00
raven score	**56.40**	**53.96**	53.02	55.49	**54.47**	54.47	0.00	**0.03**	**0.70**	0.23	0.00

Notes:

[+] p-values associated with the test that means are statistically different, taking into account clustering at the family level.

[++] pce: log of household per-capita expenditures, epe: log of extended family per-capita expenditures. Both distributions are divided in three portions, based on the cut-offs used in the main specification in the paper: percentile 25th and percentile 75th of the log of family per-capita expenditure distribution.

Source: MxFLS2 & MxFLS3

Table 28: Remittances by adult migrants interviewed in the United States

	mean	sd	median
Transfers sent to Mexico			
Percent sent transfers	**0.65**	0.48	**1.00**
Conditional on sending…			
Percent sent for consumption	**0.92**	0.27	**1.00**
Percent sent for savings/business	**0.14**	0.34	**0.00**
Amount sent	**3,229**	4,802	**1,453**
Amount sent for consumption	**2,563**	3,665	**1,240**
Amount sent for savings/business	**3,216**	3,694	**1,908**
Amount sent for consumption/total sent	**0.83**	0.32	**1.00**
Am. sent for cons/total sent (cond on sending for both)	**0.52**	0.28	**0.50**
Amount sent in cash/total sent in cash or goods	**0.83**	0.32	**1.00**
Transfers received from Mexico			
Percent received transfers	**0.09**	0.29	**0.00**
Conditional on receiving…			
Amount received	**384**	1,197	**54**
Net transfers			
Net transfers (sent - received)	**2,069**	4,176	**492**
Beneficiaries in Mexico			
Percent sent to one recipient	**0.60**	0.49	**1.00**
Percent sent to two recipients	**0.21**	0.41	**0.00**
Percent sent to three or more recipients	**0.19**	0.39	**0.00**
Last transaction			
Amount sent	**270**	359	**150**
Implied frequency of transfers	**17**	32	**10**
Cost paid for the transaction	**9**	7	**10**
Cost / total sent	**0.065**	0.073	**0.050**
Percent used a financial institution (not bank)	**0.70**	0.46	**1.00**
Percent used a bank	**0.22**	0.41	**0.00**
Percent took it herself/friend	**0.07**	0.26	**0.00**
Time in the US			
Year first arrived to the US	**2002**	7	**2004**
Percent arrived on or before 2002	**0.37**	0.48	**0.00**
Percent arrived between 2002 and 2005	**0.34**	0.47	**0.00**
Percent arrived after 2005	**0.30**	0.46	**0.00**
Expected time in the US			
Probability of living in the US in three years	**64.17**	33.58	**70**
Probability of living in the US in ten years	**45.64**	35.08	**50**
Probability of coming back to Mexico at some point	**59.47**	47.74	**50**
Expected number of years in the US before coming back	**6.80**	12.53	**3**
Other individual characteristics			
Age	**30**	11	**27**
Female	**0.40**	0.49	**0.00**
Household resources			
Log per-capita expenditures	**6.55**	0.57	**6.51**
Log per-capita income	**6.35**	0.86	**6.42**
Log per-capita wealth	**7.95**	1.73	**7.86**

Notes: All statistics on transfers correpond to transfers sent/received in the 12 months previous to the interview date.
Source: MxFLS3

Table 29: Number of pairwise tests statistically significant when analyzing household budget shares - Main specification

Family type		All families		Same locality		Spread across Mx		pread across Mx & U	
		count	weights+	count	weights+	count	weights+	count	weights+
Ratios significant at the 5% level									
household	ext family								
spl1 (0-25)	spl1 (0-25)	4	0.13	0	0.18	3	0.13	1	0.01
	spl2 (25-75)	0	0.12	0	0.15	0	0.15	2	0.06
	spl3 (75-100)	0	0.07	2	0.03	0	0.04	1	0.18
spl2 (25-75)	spl1 (0-25)	5	0.11	2	0.14	1	0.12	0	0.03
	spl2 (25-75)	0	0.18	0	0.21	0	0.20	2	0.08
	spl3 (75-100)	0	0.1	3	0.08	0	0.09	1	0.16
spl3 (75-100)	spl1 (0-25)	2	0.06	1	0.03	0	0.05	0	0.14
	spl2 (25-75)	0	0.12	0	0.09	0	0.09	0	0.19
	spl3 (75-100)	0	0.11	1	0.09	0	0.11	0	0.15
Sum		11		9		4		7	
Weighted sum		10.71		6.3		4.59		5.67	

Notes: numbers correspond to main specification used in the analysis (see Table 19)

+ % of households in the sample with household and other family resources in each spline, as defined in main specification

Source: MxFLS2 & MxFLS3

Table 30: Estimated effect of family resources on household budget shares, by family type - Linear specification

Family type	All families					Same locality *neighbor families*				
Budget share	**food** (at home & meals out)	**p. care clothing health education recreation**	**semi dur. cleaning insurance repair**	**transport comm.**	**housing** (rent & utilities)	**food** (at home & meals out)	**p. care clothing health education recreation**	**semi dur. cleaning insurance repair**	**transport comm.**	**housing** (rent & utilities)
log pce										
***household* (β1)**	-4.13	2.55	2.09	4.96	-5.48	-5.20	2.95	2.11	5.10	-4.97
	[0.538]***	[0.269]***	[0.224]***	[0.330]***	[0.419]***	[0.685]***	[0.358]***	[0.256]***	[0.383]***	[0.595]***
***other family members* (β2)**	-1.30	-0.05	0.32	-0.20	1.22	-1.41	0.02	0.11	0.02	1.26
	[0.393]***	[0.233]	[0.181]*	[0.237]	[0.330]***	[0.534]***	[0.337]	[0.223]	[0.302]	[0.464]***
Family Resources Matter										
Joint p-value (β2)			**0.00**					**0.04**		
Unitary test										
β1 - β2	-2.83	2.60	1.77	5.16	-6.70	-3.78	2.93	2.01	5.08	-6.23
p-value	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
Pareto Efficiency Test										
β1 / β2	3.18	-56.01	6.47	-24.92	-4.49	3.68	128.82	19.92	238.50	-3.94
Joint p-value+			**0.00**					**0.10**		
# observations	4,371	4,371	4,371	4,371	4,371	2,677	2,677	2,677	2,677	2,677

Family type	Spread across Mexico					Spread across Mx & US *international families*				
Budget share	**food** (at home & meals out)	**p. care clothing health education recreation**	**semi dur. cleaning insurance repair**	**transport comm.**	**housing** (rent & utilities)	**food** (at home & meals out)	**p. care clothing health education recreation**	**semi dur. cleaning insurance repair**	**transport comm.**	**housing** (rent & utilities)
log pce										
***household* (β1)**	-3.20	1.53	1.44	5.75	-5.52	-2.39	2.30	2.24	4.33	-6.48
	[1.291]**	[0.645]**	[0.561]**	[0.780]***	[0.912]***	[1.313]*	[0.554]***	[0.505]***	[0.865]***	[0.828]***
***other family members* (β2)**	-1.43	0.41	0.69	0.05	0.28	-0.93	-0.90	0.27	-0.42	1.97
	[1.118]	[0.531]	[0.538]	[0.541]	[0.667]	[0.716]	[0.419]**	[0.351]	[0.585]	[0.693]***
Family Resources Matter										
Joint p-value (β2)			**0.75**					**0.02**		
Unitary test										
β1 - β2	-1.77	1.12	0.75	5.70	-5.79	-1.46	3.20	1.97	4.74	-8.45
p-value	**0.38**	**0.19**	**0.41**	**0.00**	**0.00**	**0.31**	**0.00**	**0.01**	**0.00**	**0.00**
Pareto Efficiency Test										
β1 / β2	2.24	3.75	2.09	110.75	-19.91	2.57	-2.56	8.21	-10.40	-3.29
Joint p-value+			**0.59**					**0.10**		
# observations	756	756	756	756	756	938	938	938	938	938

Notes: Additional controls listed in notes Table 19.

+ Corresponds to last row Table 32

Source: MxFLS2 & MxFLS3

Table 31: Estimated effect of family resources on household budget shares, by family type - Non-linear specification

Family type	All families					Same locality *neighbor families*				
Budget share	**food**	**p. care**	**semi dur.**	**transport**	**housing**	**food**	**p. care**	**semi dur.**	**transport**	**housing**
log pce										
***household* (β1)**										
spline - 0050	-4.86	2.58	2.04	5.97	-5.72	-4.53	3.07	1.71	5.31	-5.56
	[0.713]***	[0.356]***	[0.238]***	[0.360]***	[0.578]***	[0.981]***	[0.470]***	[0.322]***	[0.422]***	[0.823]***
spline - 5000	-3.28	2.54	2.16	3.81	-5.22	-6.06	2.76	2.64	4.84	-4.18
	[0.919]***	[0.467]***	[0.442]***	[0.594]***	[0.619]***	[1.082]***	[0.711]***	[0.497]***	[0.742]***	[0.877]***
***other family members* (β2)**										
spline - 0050	-1.27	-0.39	0.07	-0.42	2.00	-1.24	-0.69	-0.17	0.23	1.87
	[0.608]**	[0.362]	[0.250]	[0.368]	[0.524]***	[0.797]	[0.498]	[0.317]	[0.443]	[0.697]***
spline - 5000	-1.29	0.34	0.61	-0.01	0.35	-1.68	1.10	0.53	-0.30	0.36
	[0.633]**	[0.369]	[0.324]*	[0.410]	[0.549]	[0.985]*	[0.612]*	[0.425]	[0.568]	[0.857]
Non-linearity test										
p-value β1	0.20	0.95	0.83	0.00	0.56	0.32	0.74	0.15	0.60	0.26
p-value β2	0.99	0.19	0.24	0.51	0.05	0.75	0.04	0.23	0.51	0.22
Family Resources Matter										
p-value β2	0.03	0.51	0.07	0.38	0.00	0.10	0.24	0.13	0.84	0.00
Joint p-value			**0.00**					**0.03**		
Unitary test										
β1 - β2										
spline - 0050	-3.59	2.97	1.97	6.38	-7.72	-3.30	3.76	1.88	5.08	-7.42
spline - 5000	-1.99	2.19	1.55	3.82	-5.57	-4.38	1.67	2.11	5.14	-4.54
p-value										
spline - 0050	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00
spline - 5000	0.10	0.00	0.01	0.00	0.00	0.01	0.11	0.00	0.00	0.00
Joint p-value	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
Pareto Efficiency Test										
β1 / β2										
spline - 0050	3.83	-6.59	27.57	-14.38	-2.86	3.67	-4.45	-9.87	22.78	-2.98
spline - 5000	2.55	7.40	3.56	-381.46	-15.13	3.60	2.52	5.00	-16.25	-11.71
Joint p-value^{+}			**0.02**					**0.25**		
# observations	4,371	4,371	4,371	4,371	4,371	2,677	2,677	2,677	2,677	2,677

Family type	Spread across Mexico					Spread across Mx & US *international families*				
Budget share	**food**	**p. care**	**semi dur.**	**transport**	**housing**	**food**	**p. care**	**semi dur.**	**transport**	**housing**
log pce										
household (β1)										
spline - 0050	-4.33	1.94	2.10	6.62	-6.33	-5.41	1.83	2.33	6.69	-5.43
	[1.562]***	[0.954]**	[0.627]***	[0.961]***	[1.240]***	[1.497]***	[0.741]**	[0.510]***	[0.958]***	[1.058]***
spline - 5000	-1.89	1.10	0.67	4.86	-4.74	0.21	2.70	2.17	2.34	-7.42
	[2.316]	[1.105]	[1.147]	[1.135]***	[1.408]***	[1.963]	[0.814]***	[0.867]**	[1.353]*	[1.209]***
other family members (β2)										
spline - 0050	-1.82	0.17	0.97	-0.59	1.28	-0.69	-0.50	0.41	-2.65	3.43
	[1.656]	[0.854]	[0.680]	[0.892]	[1.018]	[1.396]	[0.788]	[0.701]	[1.203]**	[1.387]**
spline - 5000	-0.82	0.66	0.28	0.74	-0.85	-0.98	-1.11	0.20	0.76	1.13
	[1.788]	[0.945]	[1.283]	[0.954]	[1.165]	[1.036]	[0.618]*	[0.496]	[0.783]	[0.963]
Non-linearity test										
p-value β1	0.41	0.61	0.34	0.22	0.41	0.02	0.43	0.88	0.01	0.21
p-value β2	0.70	0.74	0.69	0.38	0.22	0.88	0.59	0.83	0.03	0.22
Family Resources Matter										
p-value β2	0.74	0.86	0.59	0.23	0.34	0.46	0.15	0.15	0.10	0.01
Joint p-value			**0.68**					**0.01**		
Unitary test										
β1 - β2										
spline - 0050	-2.51	1.77	1.13	7.21	-7.61	-4.72	2.33	1.92	9.34	-8.86
spline - 5000	-1.07	0.44	0.39	4.12	-3.89	1.20	3.80	1.97	1.59	-8.55
p-value										
spline - 0050	0.34	0.17	0.26	0.00	0.00	0.02	0.02	0.03	0.00	0.00
spline - 5000	0.76	0.76	0.86	0.01	0.06	0.58	0.00	0.07	0.29	0.00
Joint p-value	**0.57**	**0.32**	**0.35**	**0.00**	**0.00**	**0.08**	**0.00**	**0.02**	**0.00**	**0.00**
Pareto Efficiency Test										
β1 / β2										
spline - 0050	2.38	11.66	2.17	-11.22	-4.96	7.89	-3.65	5.74	-2.52	-1.58
spline - 5000	2.30	1.67	2.40	6.59	5.58	-0.22	-2.44	10.99	3.09	-6.55
Joint p-value^{+}			**0.79**					**0.29**		
# observations	756	756	756	756	756	938	938	938	938	938

Notes: Each column represents the same budget shares used in main analysis. For additional controls see Notes in Table 19.

$^{+}$ Corresponds to last row Table 33

Source: MxFLS2 & MxFLS3

Table 32: Pareto efficiency test when analyzing household budget shares - Linear specification

Family type		**All families**	**Same locality**	**Spread across Mx**	**Spread across Mx & US**
Individual tests of pairwise ratios					
food	personal goods	0.00	0.03	0.62	0.05
	semi-durables	0.17	0.09	0.94	0.37
	transport/comm.	0.00	0.02	0.11	0.11
	housing	0.00	0.00	0.22	0.08
personal goods	semi-durables	0.17	0.78	0.69	0.02
	transport/comm.	0.83	0.98	0.43	0.22
	housing	0.04	0.04	0.36	0.66
semi-durables	transport/comm.	0.05	0.72	0.19	0.28
	housing	0.00	0.03	0.14	0.01
transport/comm.	housing	0.01	0.01	0.68	0.16
Simultaneous tests of pairwise ratios					
All ratios		**0.00**	**0.10**	**0.59**	**0.10**

Notes: p-values associated with PE test based on model in Table 30

Table 33: Pareto efficiency test when analyzing household budget shares - Non-linear specification

Family type		All families		Same locality		Spread across Mx		Spread across Mx & US	
Individual tests of pairwise ratios									
household in …		**spl1**	**spl1**	**spl1**	**spl1**	**spl1**	**spl1**	**spl1**	**spl1**
extended family in …		**spl1**	**spl2**	**spl1**	**spl2**	**spl1**	**spl2**	**spl1**	**spl2**
food	personal goods	0.01	0.29	0.01	0.48	0.45	0.71	0.33	0.21
	semi-durables	0.14	0.25	0.10	0.68	0.92	0.90	0.88	0.13
	transport/comm.	0.01	0.19	0.19	0.06	0.09	0.95	0.05	0.47
	housing	0.00	0.37	0.02	0.15	0.16	0.97	0.08	0.70
personal goods	semi-durables	0.30	0.07	0.59	0.83	0.51	0.78	0.38	0.04
	transport/comm.	0.60	0.88	0.16	0.18	0.68	0.62	0.80	0.06
	housing	0.17	0.90	0.50	0.10	0.49	0.71	0.48	0.04
semi-durables	transport/comm.	0.45	0.03	0.49	0.17	0.08	0.93	0.10	0.27
	housing	0.01	0.02	0.27	0.13	0.06	0.88	0.03	0.22
transport/comm.	housing	0.00	0.95	0.01	0.83	0.55	0.82	0.32	0.89
household in …		**spl2**	**spl2**	**spl2**	**spl2**	**spl2**	**spl2**	**spl2**	**spl2**
extended family in …		**spl1**	**spl2**	**spl1**	**spl2**	**spl1**	**spl2**	**spl1**	**spl2**
food	personal goods	0.01	0.21	0.04	0.60	0.66	0.90	0.26	0.53
	semi-durables	0.02	0.62	0.01	0.70	0.69	0.99	0.24	0.33
	transport/comm.	0.01	0.04	0.43	0.07	0.33	0.75	0.71	0.37
	housing	0.00	0.08	0.01	0.27	0.35	0.82	0.38	0.36
personal goods	semi-durables	0.65	0.51	0.77	0.51	0.67	0.95	0.61	0.13
	transport/comm.	0.99	0.46	0.25	0.08	0.72	0.59	0.05	0.14
	housing	0.03	0.20	0.10	0.09	0.78	0.63	0.05	0.28
semi-durables	transport/comm.	0.65	0.14	0.13	0.23	0.49	0.89	0.04	0.59
	housing	0.02	0.04	0.16	0.27	0.27	0.90	0.03	0.30
transport/comm.	housing	0.01	0.62	0.01	0.91	0.51	0.92	0.24	0.16
Simultaneous tests of pairwise ratios									
household in …	**extended family in …**								
spl1 (0-50)	spl1	**0.03**		**0.16**		**0.48**		**0.41**	
	spl2	**0.38**		**0.51**		**1.00**		**0.25**	
spl2 (50-00)	spl1	**0.05**		**0.10**		**0.91**		**0.16**	
	spl2	**0.41**		**0.55**		**1.00**		**0.52**	
All ratios		**0.02**		**0.25**		**0.79**		**0.29**	

Notes: p-values associated with PE test based on model in Table 31

Table 34: Estimated effect of family resources on child human capital indicators - Additional covariates

Family type	**All families**			**Same locality** *neighbor families*			**Spread across Mexico**			**Spread across Mx & US** *international families*		
Child outcomes	**height-for-age**	**years of education**	**cognitive score (%)**	**height-for-age**	**years of education**	**cognitive score (%)**	**height-for-age**	**years of education**	**cognitive score (%)**	**height-for-age**	**years of education**	**cognitive score (%)**
PANEL A												
Additional controls: household head's age, sex, years of education and marital status												
log pce												
***household* (β1)**	0.22	0.12	1.97	0.14	0.06	2.12	0.27	0.33	1.68	0.36	0.19	1.28
	[0.0556]***	[0.0519]**	[0.725]***	[0.0646]**	[0.0646]	[0.969]**	[0.154]*	[0.109]***	[1.624]	[0.136]***	[0.111]*	[1.303]
***other family members* (β2)**	0.13	0.08	1.34	0.17	0.11	1.39	0.25	0.18	3.34	0.02	0.00	0.79
	[0.0516]**	[0.0415]*	[0.621]**	[0.0650]**	[0.0529]**	[0.827]*	[0.147]*	[0.0933]*	[1.432]**	[0.0981]	[0.0729]	[1.451]
Unitary Test												
β1 - β2	0.09	0.04	0.64	-0.02	-0.05	0.73	0.02	0.15	-1.66	0.34	0.19	0.49
p-value	0.27	0.58	0.55	0.81	0.56	0.62	0.92	0.30	0.50	0.07	0.17	0.81
PANEL B												
Additional controls: household head's height, cognitive score and risk preferences												
log pce												
***household* (β1)**	0.24	0.17	2.67	0.16	0.12	2.74	0.25	0.36	1.81	0.37	0.23	1.65
	[0.0534]***	[0.0476]***	[0.693]***	[0.0640]**	[0.0623]*	[0.928]***	[0.140]*	[0.101]***	[1.458]	[0.121]***	[0.0922]**	[1.203]
***other family members* (β2)**	0.13	0.10	1.64	0.18	0.15	2.10	0.20	0.17	3.32	0.02	-0.03	0.09
	[0.0508]**	[0.0408]**	[0.612]***	[0.0636]***	[0.0528]***	[0.815]***	[0.152]	[0.0876]*	[1.456]**	[0.102]	[0.0735]	[1.384]
Unitary Test												
β1 - β2	0.11	0.07	1.03	-0.02	-0.04	0.63	0.05	0.19	-1.51	0.36	0.25	1.55
p-value	0.18	0.33	0.32	0.86	0.68	0.65	0.83	0.19	0.53	0.05	0.05	0.42

Family type	All families			Same locality *neighbor families*			Spread across Mexico			Spread across Mx & US *international families*		
Child outcomes	**height-for-age**	**years of education**	**cognitive score (%)**	**height-for-age**	**years of education**	**cognitive score (%)**	**height-for-age**	**years of education**	**cognitive score (%)**	**height-for-age**	**years of education**	**cognitive score (%)**
PANEL C												
Additional controls: log per-capita expenditures in 2002, age and education of household head in 2002												
log pce												
***household* (β1)**	0.23	0.11	2.04	0.14	0.06	2.27	0.26	0.38	2.06	0.36	0.17	1.22
	[0.0554]***	[0.0497]**	[0.755]***	[0.0674]**	[0.0643]	[1.011]**	[0.150]*	[0.106]***	[1.686]	[0.129]***	[0.101]	[1.334]
***other family members* (β2)**	0.14	0.07	1.18	0.16	0.10	1.22	0.26	0.19	3.35	0.00	0.01	1.11
	[0.0545]**	[0.0420]	[0.617]*	[0.0688]**	[0.0548]*	[0.829]	[0.152]*	[0.0954]**	[1.431]**	[0.106]	[0.0712]	[1.385]
Unitary Test												
β1 - β2	0.10	0.04	0.86	-0.03	-0.05	1.05	0.00	0.19	-1.29	0.35	0.16	0.11
p-value	0.24	0.52	0.41	0.79	0.58	0.46	0.98	0.19	0.59	0.06	0.24	0.96
PANEL D												
Additional controls: mother's height												
log pce												
***household* (β1)**	0.17	0.17	2.86	0.13	0.13	3.31	0.13	0.35	1.14	0.29	0.21	1.33
	[0.0498]***	[0.0464]***	[0.689]***	[0.0617]**	[0.0613]**	[0.950]***	[0.134]	[0.107]***	[1.409]	[0.114]**	[0.0926]**	[1.223]
***other family members* (β2)**	0.12	0.09	1.63	0.17	0.14	1.88	0.20	0.18	3.07	0.02	-0.01	0.93
	[0.0493]**	[0.0412]**	[0.621]***	[0.0619]***	[0.0530]**	[0.836]**	[0.138]	[0.0932]**	[1.405]**	[0.0961]	[0.0738]	[1.383]
Unitary Test												
β1 - β2	0.05	0.08	1.23	-0.04	-0.01	1.43	-0.07	0.16	-1.93	0.27	0.22	0.40
p-value	0.52	0.24	0.23	0.71	0.95	0.32	0.75	0.29	0.40	0.10	0.08	0.84
# observations	2,071	3,699	3,504	1,289	2,176	2,176	366	614	553	416	909	775

Notes: Controls in all specifications include age, sex, household size and number of children under 15, family size, number of children under 15 and number of female adulst in family, location (U.S. and Mexican state dummies), MxFLS2 dummy, MxFLS2 interacted with U.S. dummy, interview date.

Source: MxFLS2 & MxFLS3

Table 35: Estimated effect of family resources on child human capital indicators - Alternative samples

Family type	**All families**			**Same locality** *neighbor families*			**Spread across Mexico**			**Spread across Mx & US** *international families*		
Child outcomes	**height-for-age**	**years of education**	**cognitive score (%)**	**height-for-age**	**years of education**	**cognitive score (%)**	**height-for-age**	**years of education**	**cognitive score (%)**	**height-for-age**	**years of education**	**cognitive score (%)**
PANEL A: only children in MxFLS3												
log pce												
***household* (β1)**	0.25	0.24	3.77	0.15	0.16	3.94	0.26	0.41	1.85	0.32	0.27	4.71
	[0.0602]***	[0.0560]***	[0.860]***	[0.0718]**	[0.0777]**	[1.157]***	[0.156]*	[0.103]***	[1.670]	[0.146]**	[0.110]**	[1.848]**
***other family members* (β2)**	0.14	0.06	1.75	0.18	0.11	2.10	0.21	0.18	2.67	0.08	0.00	1.66
	[0.0567]**	[0.0518]	[0.779]**	[0.0729]**	[0.0670]*	[0.994]**	[0.154]	[0.110]*	[1.654]	[0.0963]	[0.113]	[2.126]
Unitary Test												
β1 - β2	0.11	0.17	2.02	-0.03	0.05	1.84	0.05	0.22	-0.82	0.24	0.28	3.05
p-value	0.21	0.03	0.12	0.80	0.66	0.28	0.82	0.18	0.77	0.21	0.11	0.28
# observations	1,637	2,591	2,397	964	1,451	1,452	321	505	452	352	635	493
PANEL B: only children interviewed in Mexio												
log pce												
***household* (β1)**	0.24	0.19	3.12				0.27	0.38	1.91	0.40	0.24	2.31
	[0.0543]***	[0.0490]***	[0.702]***				[0.141]*	[0.100]***	[1.429]	[0.148]***	[0.105]**	[1.263]*
***other family members* (β2)**	0.15	0.10	1.71				0.25	0.19	3.21	-0.11	-0.03	0.99
	[0.0558]***	[0.0440]**	[0.628]***				[0.147]*	[0.0947]*	[1.416]**	[0.186]	[0.0909]	[1.451]
Unitary Test												
β1 - β2	0.09	0.09	1.41				0.02	0.20	-1.30	0.50	0.27	1.32
p-value	0.32	0.21	0.18				0.91	0.19	0.58	0.06	0.08	0.51
# observations	1,900	3,544	3,504				359	602	553	252	766	775

Notes: Controls in all specifications include age, sex, household size and number of children under 15, family size, number of children under 15 and number of female adulst in family, location (U.S. and Mexican state dummies), MxFLS2 dummy, MxFLS2 interacted with U.S. dummy, interview date.
Source: MxFLS2 & MxFLS3

Table 36: Estimated effect of family resources on child human capital indicators - Non-linear specifications

Family type	**All families**			**Same locality** *neighbor families*			**Spread across Mexico**			**Spread across Mx & US** *international families*		
Child outcomes	**height-for-age**	**years of education**	**cognitive score (%)**	**height-for-age**	**years of education**	**cognitive score (%)**	**height-for-age**	**years of education**	**cognitive score (%)**	**height-for-age**	**years of education**	**cognitive score (%)**
PANEL A: non-linear effects												
log pce												
***household* (β1)**	0.24	0.18	3.11	0.16	0.13	3.36	0.27	0.38	1.90	0.39	0.24	2.32
	[0.0528]***	[0.0475]***	[0.700]***	[0.0633]**	[0.0617]**	[0.943]***	[0.139]**	[0.0989]***	[1.420]	[0.127]***	[0.0997]**	[1.284]*
***other family members* (β2)**												
spline - 0075	0.18	0.19	1.91	0.18	0.17	1.89	0.31	0.18	2.09	0.15	0.00	2.77
	[0.0682]***	[0.0607]***	[0.940]**	[0.0853]**	[0.0758]**	[1.156]	[0.182]*	[0.141]	[1.926]	[0.183]	[0.178]	[4.945]
spline - 7500	0.08	-0.04	1.38	0.23	0.09	1.99	0.15	0.20	5.90	-0.08	-0.03	0.55
	[0.103]	[0.0686]	[1.007]	[0.147]	[0.111]	[1.796]	[0.396]	[0.187]	[2.962]**	[0.156]	[0.0936]	[1.676]
Test linear effect (constant β2)												
p-value	**0.46**	**0.02**	**0.72**	**0.76**	**0.57**	**0.97**	**0.74**	**0.92**	**0.34**	**0.40**	**0.89**	**0.69**
PANEL B: interaction with household position in the pce distribution												
log pce												
***household* (β1)**	0.22	0.18	2.94	0.11	0.10	3.62	0.28	0.40	3.60	0.41	0.25	0.13
	[0.0674]***	[0.0665]***	[0.911]***	[0.0834]	[0.0844]	[1.262]***	[0.169]*	[0.128]***	[1.860]*	[0.153]***	[0.134]*	[1.634]
***other family members* (β2)**												
if household log pce<p50	0.147	0.096	1.678	0.18	0.14	1.95	0.269	0.189	3.374	0.027	-0.021	0.848
	[0.0512]***	[0.0418]**	[0.628]***	[0.0638]***	[0.0542]***	[0.839]**	[0.149]*	[0.0938]**	[1.389]**	[0.104]	[0.0746]	[1.451]
if household log pce>p50	0.154	0.100	1.747	0.20	0.16	1.86	0.266	0.180	2.726	0.019	-0.026	1.827
	[0.0521]***	[0.0428]**	[0.640]***	[0.0647]***	[0.0540]***	[0.843]**	[0.146]*	[0.0932]*	[1.446]*	[0.102]	[0.0802]	[1.525]
# observations	2,071	3,699	3,504	1,289	2,176	2,176	366	614	553	416	909	775

Notes: Additional controls include age, sex, household size and number of children under 15, family size, number of children under 15 and number of female adulst in family, location (U.S. and Mexican state dummies), MxFLS2 dummy, MxFLS2 interacted with U.S. dummy, interview date. Panel A: Results are similar if we estimate splines for the bottom 50th and top 50th percentile. There is no evidence for non-linearities in β1 (we only reject the test for cognitive scores on the sample of migrant families).

Source: MxFLS2 & MxFLS3

Table 37: Estimated effect of family resources on child human capital indicators - Alternative outcomes

Family type	**Spread across Mx & US** *international families*						
Child outcomes	**height for age**	**=1 if stunt (hfa<-2)**	**bmi for age**	**=1 if bmi<18.5**	**weight for age**	**hemoglobin level**	**=1 if hb<12**
PANEL A: linear specification							
log pce							
***household* (β1)**	0.39	-0.05	-0.11	0.03	0.06	0.11	-0.01
	[0.128]***	[0.0370]	[0.0760]	[0.0223]	[0.0666]	[0.108]	[0.0144]
***other family members* (β2)**	0.02	-0.02	0.00	0.00	-0.05	-0.17	0.03
	[0.100]	[0.0266]	[0.0635]	[0.0207]	[0.0670]	[0.104]*	[0.0152]*
Unitary Test							
β1 - β2	0.36	-0.03	-0.11	0.03	0.11	0.29	-0.04
p-value	0.05	0.52	0.31	0.39	0.25	0.06	0.06
PANEL B: non-linear specification							
log pce							
***household* (β1)**							
spline - 0050	0.41	-0.07	-0.21	0.05	0.00	-0.04	0.02
	[0.165]**	[0.0551]	[0.0946]**	[0.0278]*	[0.0823]	[0.121]	[0.0158]
spline - 5000	0.35	-0.02	0.17	-0.04	0.21	0.63	-0.10
	[0.193]*	[0.0372]	[0.154]	[0.0450]	[0.145]	[0.282]**	[0.0405]**
***other family members* (β2)**							
spline - 0050	0.32	-0.13	-0.12	-0.02	-0.01	-0.41	0.04
	[0.289]	[0.0824]	[0.197]	[0.0726]	[0.198]	[0.412]	[0.0557]
spline - 5000	-0.07	0.01	0.02	0.01	-0.06	-0.14	0.03
	[0.136]	[0.0260]	[0.0790]	[0.0238]	[0.0841]	[0.122]	[0.0195]
Test linear effect							
p-value (β1)	0.80	0.51	0.06	0.11	0.24	0.04	0.02
p-value (β2)	0.29	0.13	0.55	0.75	0.84	0.55	0.87
# observations	416	420	1,331	1,331	1,338	811	811

Notes: Additional controls in all specifications are: age, sex, household size and number of children under 15, family size, number of children under 15 and number of female adulst in family, location (U.S. and Mexican state dummies), MxFLS2 dummy, MxFLS2 interacted with U.S. dummy, interview date.

Source: MxFLS2 & MxFLS3

3. International Transfers by Mexican Migrants in the United States

3.1 Introduction

It is estimated that there are over 12 million people living in the U.S. who were born in Mexico. They are thought to represent about 30% of the U.S. foreign-born population, and account for about 10% of the entire Mexican population (Pew Hispanic Center, 2013). It is thought that many of the Mexican-born migrants retain close connections with members of their extended families in Mexico and, once they have established themselves in the U.S., they regularly send remittances back to Mexico. In 2010, it is estimated that remittances worth US$22 billion were sent from the U.S. to Mexico, which places Mexico as the third largest recipient of remittance income across the globe, behind China and India (World Bank, 2011). Unlike those countries, nearly all remittances to Mexico originate in only one country - the United States.

While the investigation of international remittances has a long history in both the scientific and policy literatures, developing a full understanding of the motivations for and the impact of these transfers has been constrained by inadequate data.[1] In particular, the high prevalence of undocumented migration and the lack of systematic records on return migration that characterize the Mexico-U.S. migration setting, make difficult to

[1] For a recent overview of the literature on remittances see Rapoport and Docquier (2006) and Ozden and Schiff (2006); for a brief review on the current state and challenges for future work see Yang (2011).

interpret evidence drawn from available data sources (Hanson, 2006).[2] In this chapter we use recently-collected and extremely rich longitudinal household survey data, which is designed to be representative of the population of recent migrants to the U.S., to provide a detailed description of cross-border transfers between Mexico and the United States.[3]

The Mexican Family Life Survey is representative of the Mexican population living in Mexico at baseline, in 2002. The study is designed to follow all baseline respondents in subsequent waves including not only those who move within Mexico but also those who move to the United States and those who subsequently return to Mexico. In principle, the sample of MxFLS respondents interviewed in the U.S. is representative of the population of all Mexicans who have moved from Mexico to the U.S. since 2002 and were living in the U.S. at the time of each follow-up survey. Preserving the representativeness of the sample is crucial if we want to learn about the transfer behavior among Mexican migrants. In Farfan et al. (2013) we describe how we interviewed 91% of the respondents believed to be in the U.S. at the time of the second round of MxFLS (in 2005-2006), and 87% of those believed to be in the U.S. at the time of the third round of MxFLS (2009-13).

---

[2] In Farfan et al. (2013) we provide direct evidence on the undercount and selectivity of the sample of recent Mexican migrants to the U.S., i.e. those who arrived to the U.S. after 2002, and are interviewed in the American Community Survey and Current Population Survey. These two data sources are widely used to estimate the stock and flows of the migrant population in the United States.

[3] For evidence on the Mexico-U.S. migration setting, see Hanson (2006), Hanson and McIntosh (2008), Chiquiar and Hanson (2005), Winters et al. (2001), Massey et al. (1994), McKenzie and Rapoport (2007), McKenzie and Rapoport (2010), Kaestner and Malamud (2014).

In this chapter we describe the economic and socio-demographic characteristics of migrants and their families who send and receive international transfers. Keeping in mind that the MxFLS sample represents relatively recent migrants is important, as we document how the demographic composition of Mexican migrants to the U.S. has changed over time. Whereas a couple of decades ago few women migrated from Mexico to the United States, that pattern has changed in recent years: 40% of adult migrants to the United States who moved since 2002 are women. Furthermore, only 19% of married males (and 1% of married females) have their spouse in Mexico. Nowadays, the modal migrant to the United Sates is not a male who leaves his family behind in Mexico and lives with other male migrants in the United States. Rather, most migrants live with their spouses and children in the United States. These changes in the composition and living arrangements of cross-border migrants likely have important implications for the nature and underlying motives for sending remittances.

Within the remittance literature, motivations to remit are broadly classified in four categories, including altruism, insurance, investment, and exchange for services.[4] While attempts have been made to disentangle the effect of each motivation separately, it remains very difficult to do so because different motivations often deliver the same empirical predictions, and migrants might have multiple motivations to remit. In this

[4] For some evidence on migrant's motivations to remit, see Amuedo-Dorantes and Pozo (2006), Amuedo-Dorantes et al. (2005), and Massey and Basem (1992) for Mexico; Cox et al. (2004), Yang (2008), and Yang and Choi (2007) for the Philippines; de la Briere et al. (2002) for Dominican Sierra; Lucas and Stark (1985) and Stark and Lucas (1988) for Botswana.

work we do not attempt to separately identify the effect of different motivations, but to use the richness of the data to describe whether empirical patterns are consistent with the alternative hypothesis posted in the literature.

One hypothesis in the literature is that migrants send money home to provide for their children. Drawing on the comprehensive set of information collected on migrant's remittances, we provide a detailed description regarding the identity of the beneficiaries of transfers in Mexico, as well as the magnitudes sent to each individual, differentiating between parents, children, siblings, other relatives, or non-relatives. When we include this information in multivariate regression models, the estimates are consistent with this hypothesis: the presence of biological children in Mexico is positively associated with both the probability of sending transfers in the last year as well as the amount sent.[5] However, as migrants lay roots in the U.S. and their immediate family members live with them in the U.S., we expect this factor to be less relevant in explaining longer-term remittance patterns. In fact, even among MxFLS migrants who do not have parents, spouse or children in Mexico, 35% sent transfers to Mexico as a means of support to the recipient.[6]

---

[5] The only exception being the probability of sending transfers on the male sample.

[6] As families reunite in the U.S. transfers across household within the U.S. can also be relevant. In the MxFLS we ask about transfers sent to and received from individuals from outside the household living in the U.S. Section 3.4 provides summary statistics on the incidence of these transfers.

It has also been suggested in the literature that migrants send transfers to Mexico to build wealth there rather than in the U.S. and this wealth is either used for retirement or to build a business to which the migrant will return. The survey instrument was designed to test this hypothesis and explicitly asks about the motivation for transfers. Respondents were able to indicate multiple motivations including to support the daily living of recipients, for savings and for investment purposes. In Section 3.4 we provide a detailed description of the stated motives of remittances, as well as on the incidence and location of savings and asset holdings. While a small percentage of migrants state sending transfers to invest or save, 17% of males and 7% of females, having savings/assets in Mexico has a significant influence on the incidence and magnitude of remittances, especially for men.

Lastly, we exploit information collected from each migrant about his/her expectations of returning someday to Mexico. As expected, those who plan to return are more likely to send transfers home.[7]

In sum, MxFLS was designed to provide new evidence on international transfers by Mexican migrants in the U.S. The combination of the panel dimension of the survey, tracking of Mexican migrants to the United States and interviewing migrants as well as their family members left behind assures these data are well-suited to provide new

[7] See McKenzie et al. (2013) for evidence on the role of expectations over future earnings on migration among Tonga migrants.

insights into the motivations for and impact of transfers in a dynamic setting and examine in greater detail both sending and receiving households. By providing an in-depth description of transfer behavior, this chapter constitutes the first step in that research agenda, as it lays the foundation for testing models that explain transfer patterns and for identifying the impact of remittances on well-being.

The structure of the chapter goes as follows: Section 3.2 describes the data, Section 3.3 presents the empirical specification, Section 3.4 describes in detail the characteristics of our migrants and their remittance behavior, Section 3.5 shows the results, and Section 3.6 concludes.

3.2 Data

The data used in the analysis is the Mexican Family Life Survey (MxFLS), an ongoing longitudinal survey that collects a rich set of information on individuals, households and communities. The first wave, conducted in 2002, includes 35,677 individuals in 8,440 households and spread out across 150 Mexican communities. At baseline, the sample is representative at the national, rural-urban and regional level. The second wave of the survey (MxFLS2) was implemented in 2005-2006, reaching a 90% re-contact rate. The third wave (MxFLS3) spans over 2009-2013 with an 87% re-contact rate.[8]

[8] This is a preliminary estimate. Intensive tracking and data cleaning are still in process, and official re-contact rates are not yet available.

MxFLS was designed to track all baseline respondents and their children born after the baseline. A feature of design that is key for this research is that MxFLS attempts to follow respondents who migrate to the United States as well as those who subsequently return to Mexico. Many studies collect information on international migrants from other household members, but few large-scale surveys have tried to follow migrants across international borders. In the second wave MxFLS interviewed 91% of those believed to be in the U.S. at the time of that survey and in MxFLS3, we have re-contacted 88% of the respondents who have moved and are living in the U.S. at the time of that survey.

The U.S. component of the survey is designed to collect comparable information with the survey conducted in Mexico, as well as to collect additional information through modules specifically designed to capture important aspects of the lives of Mexicans in the United States. Particularly relevant for this project, we designed an innovative module on transfers. Pilot work on early versions of the instrument made evident a number of challenges in collecting remittances data. For example, identification of target recipients is not straightforward since many transfers are made to one person in Mexico (to save on transfer fees) and then the funds are redistributed within Mexico. We developed a sequence of question to efficiently collect this information. We also carefully pilot-tested questions to elicit the purpose of the transfers, differentiating among transfers sent for support (for consumption or free disposal of the recipient), or transfers sent as saving or investment of the sender (for investment).

To document the migration process, and differences between males and females, we use very rich information on migration histories. In particular, in the next section we describe the characteristics around the first move from Mexico to the U.S., including whether the migrant knew someone already living in the U.S. (including how many and the specific relationship with each individual), whether the migrant came by him/herself or with others (including with how many, and the relationship with each individual if did not come alone), and the primary reason for moving to the U.S. (i.e. for work, studies, reunite with family, among others).

To complement the migrants' reports regarding whether they sent transfers for savings or investment purposes, we use information on savings and asset holdings. The MxFLS questionnaire goes over a broad list of assets, and for each category we ask whether the migrant has the corresponding asset in the U.S. or/and in Mexico. The list consists of: dwelling/land; business assets; vehicles; electronics; large appliances; small electrical appliances; savings, financial assets and other investments; other.[9]

The international migrant module is designed to provide the information needed to more fully understand the choices of migrants and their families. A module on

---

[9] For those who are old enough, in future work we will complement this information with savings/assets holdings in previous waves, especially before the migration took place. Additionally, we will use information on businesses and assets from family members who are in Mexico.

expectations about staying in the U.S. and returning to Mexico asks each migrant four questions. We begin by asking the migrant to estimate the probability that he/she will be living in the U.S. in three years and then in 10 years. We proceed to ask if they plan to ever return permanently to Mexico and, if so, how many years they expect to stay in the U.S. before returning permanently to Mexico.

Finally, we describe and use the information on living arrangements and the location of family members. Identifying household members is not straightforward in this context, as migrants usually engage in non-traditional forms of living arrangements in order to save costs. After doing some pilots we decided to define the household as "*a group of individuals who usually live together, usually consume meals provided by a common budget and usually share other expenses*" (besides housing and food). It turns out that, when following this definition, about 60% of the households share the dwelling with non-household members. However, once we identify household members in this way, the resulting household structure is very similar to that in Mexico, in that most individuals correspond to what we would call nuclear family (adult couple with children), but many households also have extended-family members.[10]

With respect to the location of close relatives, we ask about the location of each parent and a very detailed set of information on every child even born, including the age

---

[10] Among U.S. migrants, about 13% of the households are extended-family households. In Mexico that number is about 30%. Table 25 in Chapter 2 shows in detail the structure of the household roster in the U.S.

and current location, if alive. We use this information to evaluate the association there is between having close relatives in Mexico and the probability and magnitude of remittances.

3.3 Empirical specification

To estimate the incidence of transfers we run a linear probability model of the following form:

(3.1) $sent_i = \beta_0 + x_i^{'} \beta + \varepsilon_i$,

where $sent_i$ equals 1 if individual i sent transfers in the 12 months prior to the interview date, and $x_i^{'}$ is a set of covariates that vary across different specifications.

To evaluate the determinants of the amount of transfers sent, we estimate a similar regression equation of the form:

(3.2) $\log(trans)_i = \alpha_0 + z_i^{'} \alpha + \omega_i$,

where $\log(trans)_i$ is the log of the total amount sent by individual i in the 12 months prior to the interview date, and $z_i^{'}$ is the set of explanatory variables.

In both models we use the same set of covariates, starting with a baseline specification that includes basic socio-demographic characteristics, including age, education, marital status, household size, whether the household rents the dwelling, and the year to arrival to the U.S.; as well as location of interview, where the country is divided into California, Texas, Illinois, Northeast, South Midwest and West; and quarter-year interview dummies.[11]

Next we add, sequentially and one at a time, different sets of explanatory variables, which are chosen taking into account the different hypotheses posted in the migration literature. In the full specification we add all covariates simultaneously.

We first extend the baseline specification by including household and individual resources. Household resources are measured by the log of per-capita expenditure, and individual resources by labor income.[12] In both cases, we control for these factors non-parametrically by including dummy indicators of whether resources fall in the second, third or fourth quartile of the distribution.

Next we use migrants' stated expectations regarding their plans to stay in the U.S., and whether (and when) returning to Mexico. Using the information on whether the migrant expects to come back to Mexico someday, we divide the sample into those who do not expect to return ever, those who expect to come back with some positive probability, and those who report returning with certainty. Results remain the same if we use instead information collected on the expected probability of be living in the U.S. in three or ten years.

---

[11] All monetary magnitudes are expressed in 2009 US$, using PPP exchange rates when values are reported in Mexican pesos (such as the value of assets migrants have in Mexico), and annual inflation rates when the interview took place later than 2009.

[12] Results remain if we use total individual income instead of labor income, and if we control for employment status.

As mentioned in the introduction, a main motivation to send transfers to Mexico is to help support family members, especially children. In the models we separately control for whether the migrant has spouse, parents, or children in Mexico, differentiating between children 15 years old or younger, and children 16 years old or older.

Finally, we add some financial status indicators, measured by savings and asset holdings in Mexico or in the U.S.[13] We generate four indicators to differentiate among those migrants that only have savings/assets in the U.S., those who only have savings/assets in Mexico, those who have savings/assets in both countries, and those who do not have any savings/assets.

3.4 Descriptive analysis

As briefly mentioned in the introduction, the migration process differs significantly by gender. In this section we document the most important differences between our sample of men and women, which motivates and helps understand the differences we note in the models that predict the incidence and magnitude of international remittances presented in the following section.

In Table 38 we present summary statistics on a rich set of characteristics that describe different aspects of the life of migrants in the U.S., including their migration

[13] The data also has information on debt status, but we cannot separately identify debts with financial institutions from debts with family members. For that reason, we do not use this information in the models.

experience and expectations about their future in the U.S. Among others, the table includes basic socio-demographic characteristics, living arrangements, labor market outcomes, different indicators of financial status, characteristics of the first trip to the U.S., and location of immediate family members. For each outcome we present summary statistics for the whole sample, as well as for males and females separately, and show the p-value associated with the test that mean characteristics are not statistically different across the two groups.

In terms of basic socio-demographic characteristics, we see that MxFLS migrants are on average about 30 years old, with females marginally older than males. Females have on average more years of education relative to males (8.62 versus 8.17), they are more likely to be married (72% versus 46%), and they are less likely to be renting the dwelling they live in (75% versus 85%). In terms of living arrangements, they tend to live in larger households (3.77 versus 2.61), and they have more children 0 to 15 years old in the household (1.3 versus 0.56). While 36% of our male adult migrants are in a one-person household, only 7% of female adult migrants are.

Differences on key characteristics describing the first trip to the U.S., expectations over the probability of staying in the U.S. (and coming back to Mexico), and the location of children further point to the idea that females are relatively more established in the U.S. relative to males, even though they are more likely to have come in later years.

With respect to the first trip made to the U.S. to live, both males and females are equally likely to report they knew someone living in the U.S. before coming (about 76%). However, important difference emerge in the relationship between the migrant and the individuals previously living in the U.S. Virtually all females report having a relative in the U.S., while only 78% of males do so. In contrast, about 33% of males report having an acquaintance in the US, while only 5% of females do. Furthermore, not only females are more likely to have family in the U.S. before their migration takes place, but they are also more likely to come with someone in their first trip: 80% of females and 59% of males traveled with someone else. Conditional on traveling with company, 85% of females came with a relative whereas only 51% of males did. Finally, the main motive to come to live to the U.S. is quite different. Among males, the great majority reports work as the main reason to come to the U.S. (83%). In contrast, the most common reason among females is to reunite with their families (44%), followed by work (38%).

Expectations regarding how long to stay in the U.S. and when, if at all, to return to Mexico are consistent with females more likely to establish roots in the U.S. Females report a marginally higher probability of living in the U.S. in three years (65% vs 61%), and a higher probability of living in the U.S. in 10 years (48% vs 40%). Consistent with this, they also report a lower probability of living in Mexico again someday (53% versus 66%).

In terms of the location of the immediate family, females are less likely to have their close relatives in Mexico. As mentioned in the introduction, most of our migrants do not fall into the category of split-migrant households, i.e. only 34% of married males have their spouse in Mexico, and this number drops to 1% for females. Additionally, females are less likely to have their mother living in Mexico (conditional on the mother being alive), they are more likely to have children in the household (conditional on having children), and are less likely to have children living in Mexico.

Finally, we compare labor market outcomes as well as different indicators of financial status. As expected, males are significantly more likely to be employed at the time of interview (91% versus 52%), and have significantly higher labor income in the 12 months prior to the interview date (either unconditionally, or conditional on being employed). In terms of savings, on average 70% of the adult migrants have no savings, with females more likely to report zero savings relative to men. Among those with savings, we find that females are equally likely to only have savings in the U.S., and are less likely to have only savings in Mexico or savings in both countries. For both groups, between 60% and 70% of the migrants that report some savings only have savings in the U.S. Next we analyze whether our migrants have assets, and the location of those assets. About 20% of the migrants report no assets, with females more likely to report no assets relative to men. Furthermore, among those who have some assets, females are more likely to only have assets in U.S., as opposed to having assets in Mexico. Among males with some assets, 12% only has assets in the U.S., 57% has assets in Mexico and the U.S., and 31% only has assets in Mexico. The corresponding distribution for females is 5%, 62%, and 33%, respectively.

Following the same structure of Table 38, Table 39 describes migrant's behavior in terms of remittances. On average, 67% of MxFLS migrants report having sent transfers to Mexico in the last 12 months. Most transfers are destined to help the intended recipients (93% report sending transfers for this purpose), and only a few are sent with the purpose of investing or saving in Mexico (only 15% report having sent resources to save or invest). While most migrants sent transfers to Mexico, the percentage of migrants that received help from Mexico, or that sent to or received transfers from individuals within the U.S. is significantly lower (the figures being 10%, 3% and 8% respectively).

Once again, some of these statistics are significantly different between males and females. Females are significantly less likely to send transfers to Mexico (59% versus 73%).[14] When looking at the purpose of the transfers, we note that conditional on sending, females are marginally more likely to send transfers to assist individuals in Mexico, and significantly less likely to send remittances to invest or save in Mexico. Furthermore, males are not only more likely to send, but conditional on sending

---

[14] Note that this difference remains even if we construct the statistic at the household level. At the bottom of the table we present household level statistics, and we note that it is less likely that at least one person within female's households report sending transfers to Mexico relative to male's households.

transfers to Mexico they send more. On average, males sent US$4,218 in the 12 months prior to the interview date, while females sent on average US$2,076. These magnitudes account for about 47% of household per-capita expenditure among males, and 26% of household per-capita expenditure among females. If we also take into account the value of transfers received, we note that males sent on average a net value of US$3,047, and females sent US$1,185. In terms of transfers sent to or received from other households in the U.S. there are no significant differences between the two groups.

Exploiting the more detailed information collected on the last transaction made by each migrant, we can infer the implied frequency of the transactions as well as the average cost migrants pay to send transfers to Mexico. Using information on the total amount sent in the last 12 months, and the amount sent on the last transaction, we note that males make on average more transactions and send more resources per transaction relative to females. Males sent on average US$320 per transaction, and transferred 21 times in the last year. In turn, females transferred on average US$221 per transaction, and transferred about 14 times in the last year. In terms of how migrants transfer the resources, most migrants use a financial institution that is not a bank, like Western Union or the like (69%), a smaller group uses a bank (23%), and the reminder carried the money themselves or sent it with friends or family. Finally, in terms of costs, it seems the cost per transaction is pretty much determined (US$10), which means that the cost as a percentage of the amount sent is largely determined by the amount and frequency with which migrants send transfers to Mexico. Given that males send more resources, but they also send more frequent, it is not clear a priori whether they pay more per transaction. As it turns out, the average cost paid is lower among males than among females, 6% and 8% respectively.

Finally, we describe the number and identity of the beneficiaries of the transfers in Mexico. On average, about 40% of the migrants sent transfers to more than one recipient in Mexico, about half of which sent to two beneficiaries and the other half sent to three or more. While the distribution of the number of recipients is not different between males and females, the relationship with the beneficiaries changes between the two groups. Conditional on having a mother in Mexico, females are more likely to send transfers to their mother (80% versus 67%), though there is no difference when it comes to sending to the father (about 30% among both groups). Conditional on having children in Mexico, females are also more likely than men to send transfers directly to their children. This last effect remains if we condition on having children 16 years old or older, which is important if we consider that transfers to younger children are most likely made through transfers to parents or other relatives. In terms of less close relatives, we have that females are ten percentage points more likely to send transfers to siblings (28% versus 18%), but equally likely to send transfers to other relatives (about 17%) and to non-relatives (about 6%).

3.5 Results

Before going into the models, Figure 6 presents basic non-parametric relationships between the probability of sending transfers and basic individual characteristics. Looking at the age profile we note that relative to individuals 21 to 25 years old, it is only younger migrants (15 to 20) who are significantly less likely to send transfers. There is no statistically significant difference between the baseline group and older migrants, and these results are similar for both males and females. With respect to education levels, it is only among males that we see differences in the incidence of transfers. In particular, male migrants with high-school complete or more are significantly less likely to send transfers relative to migrants with lower education. When we look at the relationship between migrant's expectations over the probability of returning to Mexico and the incidence of transfers, we see that there is a clear positive association between these two variables. Relative to migrants that state they are returning to Mexico with zero probability, those who expect to come back with some probability are more likely to send transfers, and those who expect to return with certainty are even more likely to send transfers to Mexico. This is consistent with migrants keeping close links with Mexico if they expect to come back at some point. Lastly we present the relationship between transfers and household resources, measured by the log of per-capita expenditures. Among males, migrants in the second and third quartile of the distribution are marginally less likely to send transfers relative to migrants at the bottom of the distribution, while migrants in the fourth quartile are more likely to send transfers to Mexico. Among females the pattern is different. There is a positive association between the quartile in the distribution and the probability of sending transfers, with females in the third and fourth quartiles significantly more likely to send transfers to Mexico.

Next we present the results of the two models specified in Section 3.3. For both males and females, we run a number of specifications to predict the incidence of transfers, and use the same models to predict the amount sent conditional on sending remittances to Mexico. In this way we can easily compare the effects of each set of explanatory variables on both outcomes. Results are shown in Table 40 for males and Table 41 for females. At the bottom of each table we present the p-value associated with the test that the joint effect of each group of covariates is equal to zero. Finally, we run a fully interactive model that allows us to see whether the differences between males and females are statistically significant. The coefficients and standard errors associated with each interaction term are shown in Table 42.

At this time we do not differentiate between motives. As mentioned in previous section, the vast majority of transfers are meant to be used by the recipient, and only a small number of migrants reported sending some transfers specifically to invest or save.

3.5.1 Results for males

The first column in Table 40 presents the results that correspond to the baseline specification, which controls for basic socio-demographic characteristics, including age, education, marital status, household size, whether the household rents the dwelling, and the year to arrival to the U.S.[15] Similarly to what we saw in Figure 6, younger migrants (age 15 to 20) are significantly less likely to send transfers relative to the omitted category of migrants 26 to 30, and migrants who arrived to the U.S. in later years are more likely to send. In contrast, there is no significant relationship between the incidence of transfers and the migrant's level of education.

In terms of the other set of covariates, we see that marital status is only marginally significant in this model, but household characteristics are significant predictors. Migrants renting their dwelling are more likely to send transfers home, and there is a negative association between household size and incidence of transfers. Relative to migrants in households with 2 to 4 individuals, migrants in one-person household (migrants in larger households) are significantly more likely (less likely) to send transfers home. This is consistent with the idea that migrants in larger households are more likely to have reunited with their immediate family, and therefore less likely to remain connected with Mexico.

[15] Remember all models also control for interview date and location of interview.

Columns (2) to (6) include additional sets of covariates one at a time, and column (7) adds them all. In columns (2) and (3) we control for household and individual resources, in column (4) we control for the migrant's expectations regarding the probability of returning to Mexico, in column (5) we control for the location of immediate relatives, and finally in column (6) we add some indicators regarding the financial status of the migrant (savings and assets holdings). Looking across columns we note that the two variables that remain significant in the full specification are renting and year of arrival to the U.S.

Contrary to what we saw in Figure 6, after controlling for a baseline set of covariates, there isn't a significant association between the incidence of transfers and household resources measured by the log of per-capita expenditures. In column (3) we further control for individual resources, measured by the log of labor income. As it turns out, individual income does have a marginally significant effect on the probability of sending transfers. Individuals in the third and fourth quartile of the distribution are more likely to send transfers home relative to individuals in the first quartile. As evident in column (7) the effect of individual resources disappears once we control for the location of close relatives and financial status indicators.

In column (4) we can look at the effect of migrant's expectations regarding the probability of returning to Mexico. As noted in Figure 6, relative to migrants that do not expect to come back ever, those who do are significantly more likely to send transfers home.

In column (5) we use the presence of close relatives in Mexico as indicators of a stronger connection of our migrants with their home country. While those who have parents in Mexico are significantly more likely to send transfers home, after we control for the presence of parents in Mexico, neither having a spouse nor having children affects the incidence of transfers. As pointed out below, this is one of the differences between males and females, as having young children does increase the probability of sending transfers home among females.

Lastly, we analyze whether migrants with savings or assets in Mexico are more likely to send remittances. For each of these two outcomes, we differentiate between those who only have savings/assets in Mexico, those who only have savings/assets in the U.S., and those who have savings/assets in both countries. In both cases, those who only have savings/assets in the U.S. are not more or less likely to send transfers to Mexico relative to those with no savings/ assets, but migrants with assets/savings in Mexico (either only in Mexico or in both countries) are more likely to send remittances.

In the last column we run a model that incorporates the covariates from all the models simultaneously. The results suggest that when following this full specification, a few factors remain significant in predicting the incidence of remittance, including the time they've been in the U.S., the expectations regarding returning to Mexico, having parents in Mexico, and having savings or assets in Mexico.

Next we analyze the effect of the same set of variables on the amount of remittances sent in the 12 months prior to the interview date. In the right panel of Table 40 we present the same models used to predict the incidence of transfers, and note some interesting differences in their effect when it comes to predict the amount sent. Education and household resources continue to have no explanatory power in predicting remittance behavior, but individual labor income is significant in these models. Migrants with higher labor income send more remittances to Mexico. Another difference is that having spouse and children in Mexico are associated with higher remittances.

The remaining variables, i.e. household composition, year of arrival, expected probability of returning to Mexico, and having savings/assets in Mexico, have an impact that goes in the same direction as that we found on the probability of sending remittances to Mexico.

3.5.2 Results for females

In Table 41 we present the results for females. Among the covariates in the baseline specification, we see that it is only marital status and household size the characteristics that are significantly associated with the incidence of transfers: married females are more likely to send transfers, and females in larger households are less likely to send transfers.

The effect of household resources is also different relative to what we saw for males. In this case, women in households with higher per-capita expenditures are more likely to send transfers home. However, once we control for individual labor income, household resources are no longer significant.

Similar to the case of males, females who expect to return to Mexico someday are more likely to send transfers relative to those who do not expect to come back. However, this effect also disappears once we control for the full set of explanatory variables (column (7)).

With respect to the location of close relatives, in this case we only analyze the effect of having parents or children in Mexico, because only a handful of females in the sample have a spouse in Mexico. Female migrants that have their mother or father in Mexico are significantly more likely to send transfers to Mexico, as do female migrants with young children in Mexico. As mentioned before, the significant positive effect of having young children on the incidence of remittances is one of the differences between males and females, a difference that is also confirmed in the fully interacted model presented in Table 42.

Finally we look at the effect of savings and assets on the probability of sending remittances to Mexico. In this case, the saving indicators have no explanatory power on this model. This is consistent with the descriptive analysis presented in the previous section, where we noted that females are significantly less likely to send transfers for investment or saving purposes relative to men (only 8% did), and only a very small share of females has any savings in Mexico. Having assets in Mexico has a marginal effect on the incidence of transfers, but the effect disappears in the full specification.

In column (7) we present the results of the full specification. For the case of females, we see that it is marital status, individual labor income, and the presence of close relatives in Mexico, the factors that remain significant in predicting who sends transfers back home.

Lastly, the right panel of the table presents the effect that these explanatory variables have on the amount of transfers sent. In all specifications we see a positive relationship between age and amount sent. More specifically, women 25 years old or younger send significantly less transfers relative to older migrants. In contrast to what we saw for males, these age effects remain after controlling for financial status, which is consistent with the evidence that saving or asset holdings are not relevant in explaining transfer behavior among females. Similarly, neither household nor individual resources are significant in explaining how much female migrants send to Mexico. Three factors turn out relevant in this model: married women send significantly more resources to Mexico relative to single women; women in one-person households send significantly more compared with women in larger households; women in the top quartile of the labor income distribution send significantly more relative to women at the bottom of the distribution; and women with parents or children in Mexico send significantly more than women without such relatives.

In Table 42 we show the coefficient on the interaction between each explanatory variable and a female dummy, when running a fully interactive model. That table reinforces the main differences between the models for males and females highlighted above. When predicting who is more likely to send transfers to Mexico, being married significantly increases the probability for females but has no effect on males. In contrast, renting a dwelling increases the incidence among males but is not relevant for females. Similarly, high levels of individual labor income, and having young children in Mexico, have a significant positive effect on the probability of sending transfers home among females, but not among males. With respect to the amount of transfers sent to Mexico, the main differences are also related to the effects of marital status and having young children in Mexico, as well as asset holdings.

3.6 Conclusion

This research exploits the richness of data on transfers collected in the Mexican Family Life Survey to carefully describe the ways resources are transferred, the costs of transfers and the ways that resources are redistributed in Mexico after they have been transferred. We use an extensive set of information collected from Mexican migrants interviewed in the U.S. to describe in a multivariate regression framework the incidence and magnitude of cross-border transfers.

We describe first the migrant population, in particular the difference between males and females, as that informs on the difference we find in transfer behavior across genders. In general, the evidence is consistent with females being more likely to lay roots in the U.S. relative to males. For example, they are more likely to come to live to the U.S. to reunite with their family, they are less likely to have their spouse in Mexico, less likely to have children in Mexico and more likely to have children in the U.S., less likely to return to Mexico someday, and more likely to stay in the U.S. for longer (conditional on ever coming back).

Consistent with the literature, we also find evidence of migrant sending substantial remittances back home. On average, MxFLS migrants send the equivalent of about 40% of household per-capita expenditures. Also in line with previous evidence, migrants tend to make frequent transactions a year. On average, the implied frequency in our sample is about 18 transactions in 12 months. This results in relatively high transactions costs, with an average cost of about 7%.

We then analyze the determinants of the incidence and magnitude of cross-border remittances. When running a full specification, we find that males who have their parents in Mexico, those who expect to return to Mexico with certainty, and those who have assets or savings in Mexico are significantly more likely to send remittances. The same variables are statistically significant in predicting how much these migrants sent in the last year, as it is having spouse and children in Mexico. In contrast to males, labor income and having spouse or children in Mexico are associated with a higher incidence of transfers among females, while expectations regarding returning to Mexico, and assets or savings holdings, do not. Conditional on sending transfers, it seems that having relatives in Mexico is the main factor in explaining how much to send.

Building on this evidence, future work will incorporate information collected in Mexico from remittance recipients, and include additional details on other family members, including the whereabouts of others who have moved to the U.S., to test hypotheses about the ways in which resources are shared among extended families. This analysis will complement the research presented in Chapter 2 regarding resource allocation within extended families. Finally, we will adapt the identification strategy implemented in Chapter 1, which is based on the biology of child growth, to provide new evidence on the impact that migration and remittances have on child nutrition.

3.7 Primary figures

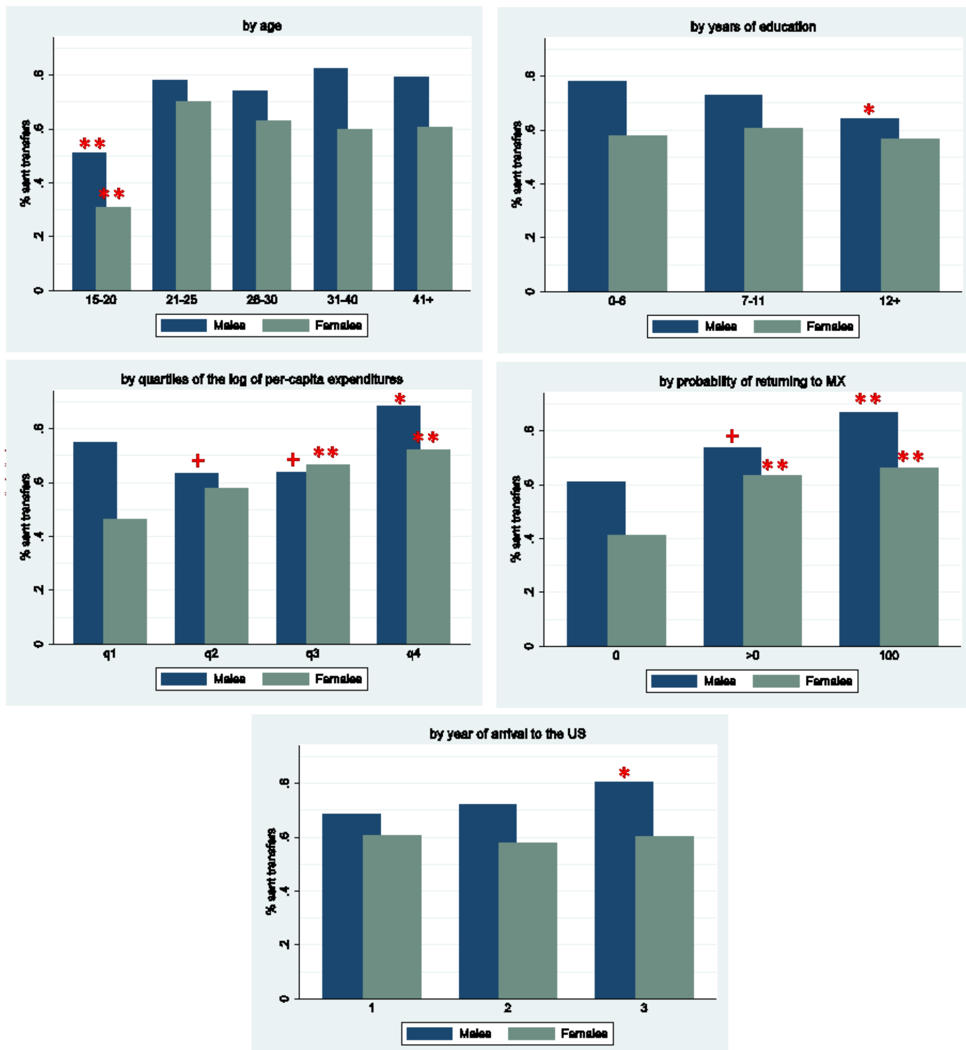


Figure 6: Incidence of transfers sent to Mexico by Mexican migrants in the United States

Source: MxFLS3 - **Significant at 1%,*Significant at 5%,+Significant at 10%.

3.8 Primary tables

Table 38: Descriptive statistics of adult migrants interviewed in the United States, by gender

	All			Male			Female			Difference[+]
	mean	sd	median	mean	sd	median	mean	sd	median	(p-value)
Individual characteristics										
Age	29.95	11.76	27	29.39	11.80	25	30.74	11.67	28	0.09
Female	0.41	0.49	0							
Years of education	8.36	3.33	9	8.17	3.32	9	8.62	3.33	9	0.05
primary	0.32	0.47	0	0.33	0.47	0	0.31	0.46	0	0.32
some high	0.48	0.50	0	0.49	0.50	0	0.47	0.50	0	0.48
high or more	0.20	0.40	0	0.18	0.38	0	0.22	0.41	0	0.20
Married	0.56	0.50	1	0.46	0.50	0	0.72	0.45	1	0.00
Born in rural locality	0.77	0.42	1	0.77	0.42	1	0.77	0.42	1	0.92
Labor market										
Employed	0.75	0.43	1	0.91	0.28	1	0.52	0.50	1	0.00
Employee	0.95	0.23	1	0.97	0.17	1	0.88	0.32	1	0.00
Self-employed	0.04	0.21	0	0.02	0.15	0	0.10	0.30	0	0.00
Labor income										
Unconditional	14,438	14,273	12,600	16,810	15,853	15,000	9,368	8,053	7,360	0.00
Conditional on being employed	15,141	14,476	13,440	17,062	15,907	15,000	10,335	8,309	9,600	0.00
Non-labor income	555	8636	0	644	11105	0	423	1579	0	0.71

	All			Male			Female			Difference+
	mean	sd	median	mean	sd	median	mean	sd	median	(p-value)
Financial status										
Has savings in the US	0.23	0.42	0	0.25	0.43	0	0.21	0.40	0	0.12
Savings in the US	3,303	5,697	1,500	2,761	3,424	1,400	4,230	8,207	1,800	0.12
Has savings in Mexico	0.12	0.33	0	0.15	0.36	0	0.08	0.27	0	0.00
Savings in Mexico	3,849	4,942	2,400	3,797	4,304	2,400	4,030	6,840	2,400	0.85
Does not have savings	0.69	0.46	1	0.65	0.48	1	0.75	0.44	1	0.00
Only has savings in Mexico	0.08	0.27	0	0.10	0.30	0	0.05	0.22	0	0.01
Only has savings in the US	0.19	0.39	0	0.20	0.40	0	0.18	0.38	0	0.52
Has savings in Mexico and in the US	0.04	0.20	0	0.05	0.23	0	0.03	0.16	0	0.05
Has assets in the US	0.73	0.45	1	0.73	0.45	1	0.72	0.45	1	0.91
Wealth in the US	10,343	40,762	2,650	8,198	20,428	3,000	13,445	58,757	2,000	0.11
Has assets in Mexico	0.39	0.49	0	0.45	0.50	0	0.30	0.46	0	0.00
Wealth in Mexico	13,142	24,800	6,400	12,899	20,890	7,000	13,661	31,678	3,960	0.79
Does not have assets										
Only has assets in Mexico	0.21	0.41	0	0.20	0.40	0	0.24	0.43	0	0.10
Only has assets in the US	0.06	0.24	0	0.08	0.27	0	0.04	0.18	0	0.01
Has assets in Mexico and in the US	0.40	0.49	0	0.36	0.48	0	0.46	0.50	0	0.00
Household characteristics										
Number of individuals in dweling	4.98	2.59	5	4.76	2.60	4	5.29	2.56	5	0.01
Number of relatives in dweling	4.01	2.34	4	3.63	2.49	4	4.55	1.99	4	0.00
Household size	3.09	1.75	3	2.61	1.76	2	3.77	1.49	4	0.00
Percent w/ household size =1	0.24	0.43	0	0.36	0.48	0	0.07	0.25	0	0.00
Number children 0-15 years old	0.87	1.06	0	0.56	0.90	0	1.30	1.11	1	0.00
Renting	0.81	0.39	1	0.85	0.36	1	0.75	0.43	1	0.00
Per-capita expenditures	846	802	650	963	659	775	686	942	549	0.00

	All			Male			Female			Difference+
	mean	sd	median	mean	sd	median	mean	sd	median	(p-value)
Time in the United States										
Year first arrived to the US	2003	6	2004	2002	7	2004	2004	6	2004	0.00
Arrived on or before 2002	0.30	0.46	0	0.34	0.47	0	0.24	0.43	0	0.00
Arrived between 02 & 05	0.37	0.48	0	0.35	0.48	0	0.42	0.49	0	0.04
Arrived after 2005	0.33	0.47	0	0.31	0.46	0	0.34	0.48	0	0.37
Fist trip to the United States										
Knew someone already living in the US	0.76	0.43	1	0.74	0.44	1	0.80	0.40	1	0.36
Had a relative (cond on knowing someone)	0.85	0.36	1	0.78	0.42	1	1.00	0.00	1	0.00
Had acquaintances (cond on knowing someone)	0.24	0.43	0	0.33	0.47	0	0.05	0.22	0	0.00
Traveled alone	0.34	0.48	0	0.41	0.49	0	0.20	0.40	0	0.01
Traveled with relatives (cond on trav w/someone)	0.64	0.48	1	0.51	0.50	1	0.85	0.36	1	0.00
Traveled with other (cond on trav w/someone)	0.36	0.48	0	0.48	0.50	0	0.17	0.38	0	0.00
Main motive when coming to the US										
Work	0.69	0.46	1	0.83	0.38	1	0.38	0.49	0	0.00
Family reasons	0.18	0.38	0	0.06	0.23	0	0.44	0.50	0	0.00
Other	0.11	0.32	0	0.10	0.30	0	0.15	0.36	0	0.37
First trip before 2002	0.16	0.37	0	0.20	0.40	0	0.11	0.31	0	0.00
Expected time in the US										
Probability of living in the US in 3 years	62.48	34.08	70	60.70	34.71	70	64.94	33.09	75	0.09
Percent report low (below 50%)	0.27	0.44	0	0.28	0.45	0	0.25	0.43	0	0.28
Percent report high (between 50 & 99%)	0.44	0.50	0	0.43	0.50	0	0.46	0.50	0	0.47
Percent report with certainty (100%)	0.29	0.45	0	0.28	0.45	0	0.29	0.46	0	0.79
Probability of living in the US in 10 years	43.24	34.68	50	39.65	34.38	50	48.20	34.53	50	0.00
Percent report low (below 50%)	0.47	0.50	0	0.50	0.50	0	0.42	0.50	0	0.05
Percent report high (between 50 & 99%)	0.40	0.49	0	0.39	0.49	0	0.41	0.49	0	0.58
Percent report with certainty (100%)	0.14	0.34	0	0.11	0.32	0	0.17	0.37	0	0.04
Probability of coming back to Mexico at some point	60.37	36.71	60	65.99	35.25	80	52.63	37.31	50	0.00
Percent report 0	0.12	0.32	0	0.09	0.29	0	0.15	0.35	0	0.03
Percent report positive (between 1 & 99%)	0.56	0.50	1	0.53	0.50	1	0.60	0.49	1	0.07
Percent report with certainty (100%)	0.31	0.46	0	0.37	0.48	0	0.24	0.43	0	0.00
Expected # of years in the US before coming back	6.37	11.79	3	5.55	10.14	3	7.61	13.85	4	0.05

	All			Male			Female			Difference+
	mean	sd	median	mean	sd	median	mean	sd	median	(p-value)
Relatives in Mexico										
Father is alive	0.82	0.38	1	0.84	0.36	1	0.79	0.41	1	0.07
Mother is alive	0.86	0.34	1	0.87	0.34	1	0.85	0.35	1	0.57
Father in Mexico (cond on being alive)	0.72	0.45	1	0.74	0.44	1	0.69	0.46	1	0.22
Mother in Mexico (cond on being alive)	0.76	0.43	1	0.82	0.39	1	0.69	0.46	1	0.00
Mother and father in Mx (cond on both alive)	0.65	0.48	1	0.69	0.46	1	0.59	0.49	1	0.01
Spouse in Mexico (cond on having a spouse)	0.17	0.37	0	0.34	0.47	0	0.01	0.09	0	0.00
Has children alive	0.58	0.49	1	0.47	0.50	0	0.75	0.43	1	0.00
Has children in household (cond on having ch alive)	0.71	0.45	1	0.52	0.50	1	0.89	0.31	1	0.00
# of children in hh	1.38	1.18	1	0.96	1.15	1	1.75	1.09	2	0.00
# of children in hh (cond on having ch in hh)	1.93	0.95	2	1.85	0.94	2	1.96	0.96	2	0.28
Has children outside the household	0.45	0.50	0	0.63	0.48	1	0.28	0.45	0	0.00
Has children in Mexico	0.32	0.47	0	0.48	0.50	0	0.17	0.38	0	0.00
# of children in Mexico	0.71	1.34	0	1.11	1.55	0	0.35	1.00	0	0.00
# of children in Mx (cond on having ch in Mx)	2.25	1.50	2	2.32	1.50	2	2.06	1.52	1	0.33
Has children […] in household										
0-5 years old	0.45	0.50	0	0.30	0.46	0	0.57	0.50	1	0.00
6-10 years old	0.19	0.39	0	0.11	0.31	0	0.26	0.44	0	0.00
11-15 years old	0.12	0.33	0	0.07	0.26	0	0.17	0.37	0	0.00
16-18 years old	0.16	0.37	0	0.13	0.33	0	0.19	0.39	0	0.04
18 and older	0.11	0.31	0	0.10	0.30	0	0.12	0.32	0	0.59
Has children […] in Mexico										
0-5 years old	0.08	0.27	0	0.15	0.36	0	0.01	0.11	0	0.00
6-10 years old	0.11	0.31	0	0.19	0.39	0	0.04	0.20	0	0.00
11-15 years old	0.09	0.29	0	0.14	0.35	0	0.05	0.22	0	0.00
16-18 years old	0.13	0.34	0	0.18	0.39	0	0.09	0.28	0	0.00
18 and older	0.15	0.36	0	0.20	0.40	0	0.10	0.31	0	0.00

Notes: Sample size: 900 (531 males and 369 females). + Test difference in means between males and females.

Source: MxFLS3

Table 39: Descriptive statistics of transfers sent by adult migrants interviewed in the United States, by gender

	All			Male			Female			Difference+
	mean	sd	median	mean	sd	median	mean	sd	median	(p-value)
Transfers sent to Mexico										
Sent transfers	0.67	0.47	1	0.73	0.44	1	0.59	0.49	1	0.00
Conditional on sending…										
Sent for consumption	0.93	0.26	1	0.92	0.28	1	0.95	0.21	1	0.08
Sent for savings/business	0.15	0.35	0	0.19	0.39	0	0.07	0.25	0	0.00
Amount sent	3,443	5,014	1,555	4,218	5,232	2,480	2,076	4,287	689	0.00
Amount sent for consumption	2,781	3,949	1,377	3,412	4,353	1,968	1,713	2,855	668	0.00
Amount sent for savings/business	3,385	3,872	1,908	3,496	3,517	1,968	2,857	5,369	858	0.56
Amount sent for consumption/total sent	0.84	0.31	1	0.81	0.33	1	0.89	0.26	1	0.00
Am sent for cons/total sent (cond on sending for both)	0.51	0.28	0.48837	0.49	0.26	0.45	0.62	0.33	0.73	0.10
Amount sent in cash/total sent in cash or in kind	0.84	0.31	1	0.86	0.30	1	0.80	0.33	1	0.03
Amount sent/labor income	0.36	0.75	0.18	0.38	0.84	0.20	0.31	0.47	0.15	0.40
Amount sent/total income	0.35	0.75	0.17	0.37	0.84	0.20	0.30	0.49	0.14	0.36
Amount sent/per-capita expenditure	0.40	0.64	0.19	0.47	0.69	0.27	0.26	0.53	0.12	0.00
Transfers received from Mexico										
Received transfers	0.10	0.30	0	0.10	0.29	0	0.10	0.30	0	0.73
Conditional on receiving…										
Amount received	391	1282	49	334	948	49	470	1649	57	0.64

	All			Male			Female			Difference+
	mean	**sd**	**median**	**mean**	**sd**	**median**	**mean**	**sd**	**median**	**(p-value)**
Net transfers										
Net transfers (sent - received)	2,280	4,428	492	3,047	4,839	1,097	1,185	3,490	191	0.00
Did not sent or received transfers from Mexico	0.30	0.46	0	0.24	0.43	0	0.37	0.48	0	0.00
Did not sent trans / Did receive transfers from Mx	0.03	0.17	0	0.02	0.15	0	0.04	0.19	0	0.18
Sent transfers / Did not receive transfers from Mx	0.60	0.49	1	0.66	0.48	1	0.52	0.50	1	0.00
Sent and received transfers from Mexico	0.07	0.26	0	0.07	0.26	0	0.07	0.25	0	0.63
Transfers sent to/received from other households in the US										
Sent transfers	0.03	0.17	0	0.03	0.17	0	0.03	0.18	0	0.71
Amount sent	1033	1541	394	1118	1845	429	923	1118	394	0.76
Received transfers	0.08	0.26	0	0.06	0.25	0	0.09	0.29	0	0.12
Amount received	431	1,214	95	569	1,660	95	294	446	97	0.36
Last transaction										
Amount sent	284	377	160	320	423	200	221	270	100	0.00
Implied frequency of transfers	18	34	10	21	31	12	14	39	6	0.04
Cost paid for the transaction	9.23	6.69	10	9.42	4.85	10	8.91	9.07	10	0.41
Cost/total sent	0.07	0.08	0.05	0.06	0.07	0.05	0.08	0.08	0.05	0.03
Used a financial institution (not bank)	0.69	0.46	1	0.73	0.45	1	0.63	0.48	1	0.02
Used a bank	0.23	0.42	0	0.21	0.40	0	0.27	0.45	0	0.08
Through a friend or him/herself	0.07	0.25	0	0.05	0.22	0	0.09	0.29	0	0.09

	All			Male			Female			Difference+
	mean	sd	median	mean	sd	median	mean	sd	median	(p-value)
Beneficiaries in Mexico										
Sent to one recipient	0.61	0.49	1	0.64	0.48	1	0.56	0.50	1	0.10
Sent to two recipients	0.21	0.41	0	0.20	0.40	0	0.22	0.42	0	0.50
Sent to three or more recipients	0.18	0.39	0	0.17	0.37	0	0.21	0.41	0	0.17
Conditional on having […] in Mexico										
Sent transfers to spouse	0.87	0.34	1	0.86	0.35	1				
Sent transfers to mother	0.72	0.45	1	0.67	0.47	1	0.80	0.40	1	0.01
Sent transfers to father	0.31	0.46	0	0.31	0.46	0	0.33	0.47	0	0.72
Sent transfers to son/daughter	0.36	0.48	0	0.24	0.43	0	0.69	0.47	1	0.00
Sent trans to son/daughter (cond on ch 16+)	0.39	0.49	0	0.26	0.44	0	0.71	0.47	1	0.00
Unconditional										
Sent to sibling	0.22	0.41	0	0.18	0.39	0	0.28	0.45	0	0.01
Sent to other relatives	0.17	0.38	0	0.17	0.37	0	0.18	0.39	0	0.62
Sent to non-relatives	0.05	0.22	0	0.05	0.21	0	0.06	0.24	0	0.41
Amount sent to spouse	5,428	5,485	4,522	5,494	5,507	4,578				
Amount sent to mother	2,326	2,802	1,211	3,065	3,132	1,968	1,232	1,731	572	0.00
Amount sent to father	2,764	3,579	1,254	3,608	4,062	2,131	1,397	2,000	689	0.00
Conditional on sending to more than one recipient…										
Amount sent to son/daughter	4,619	9,642	2,027	2,731	3,361	1,476	6,588	13,214	3,129	0.17
Amount sent to sibling	2,468	4,204	1,004	2,597	3,965	1,453	2,322	4,493	689	0.73
Amount sent to other relatives	2,812	4,890	1,049	3,605	5,694	1,476	1,621	3,052	322	0.07
Amount sent to non-relatives	3,728	5,673	1,486	5,216	6,681	3,345	1,699	3,181	443	0.12
Conditional on having […] in Mexico										
Amount sent to spouse/total sent	0.85	0.33	1	0.85	0.34	1				
Amount sent to mother/total sent	0.66	0.45	1	0.63	0.47	1	0.72	0.42	1	0.06
Amount sent to father/total sent	0.27	0.43	0	0.27	0.43	0	0.27	0.43	0	1.00
Am sent to mother/total sent (cond on mo & fa in Mx)	0.63	0.46	1	0.61	0.47	1	0.69	0.44	1	0.14
Am sent to father/total sent (cond on mo & fa in Mx)	0.25	0.42	0	0.26	0.43	0	0.22	0.40	0	0.40
Amount sent to son/daughter	0.27	0.43	0	0.15	0.34	0	0.60	0.48	0.99	0.00
Amount sent to son/daughter (cond on ch 16+)	0.31	0.46	0	0.17	0.37	0	0.67	0.49	1	0.00

	All			Male			Female			Difference+
	mean	**sd**	**median**	**mean**	**sd**	**median**	**mean**	**sd**	**median**	**(p-value)**
Household statistics										
Someone in the household sent transfers to Mexico	0.76	0.43	1	0.79	0.41	1	0.73	0.45	1	0.05
Someone in the hh received transfers from Mexico	0.23	0.42	0	0.22	0.41	0	0.25	0.43	0	0.36
Did not sent or received transfers from Mexico	0.21	0.41	0	0.18	0.39	0	0.24	0.43	0	0.03
Did not sent but did receive transfers from Mexico	0.03	0.17	0	0.03	0.18	0	0.03	0.17	0	0.85
Sent but did not received transfers from Mexico	0.56	0.50	1	0.60	0.49	1	0.51	0.50	1	0.01
Sent and received transfers from Mexico	0.20	0.40	0	0.19	0.39	0	0.22	0.41	0	0.29
Net transers (cond on having sent or received transfers)	2,875	3,622	1,771	3,202	3,711	1,968	2,374	3,426	1,377	0.00

Notes: Sample size: 900 (531 males and 369 females). + Test difference in means between males and females.

Source: MxFLS3

Table 40: Transfers sent to Mexico in the last year by male migrants

	Sent transfers							Log amount sent						
	base	hh res	ind res	pback	relatives	financial	all	base	hh res	ind res	pback	relatives	financial	all
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Age														
Base: 26-30 years old														
15-20 years old	-0.208	-0.213	-0.15	-0.209	-0.154	-0.189	-0.0982	-0.00973	-0.0496	0.0316	-0.00419	0.119	-0.147	-0.0177
	[0.0642]***	[0.0647]***	[0.0643]**	[0.0640]***	[0.0663]**	[0.0636]***	[0.0656]	[0.278]	[0.280]	[0.273]	[0.277]	[0.283]	[0.274]	[0.265]
21-25 years old	0.0034	0.0016	-0.00225	0.00462	0.00404	0.000642	0.00309	0.115	0.0928	0.0589	0.145	0.173	0.0656	0.0815
	[0.0524]	[0.0525]	[0.0514]	[0.0522]	[0.0525]	[0.0511]	[0.0506]	[0.207]	[0.208]	[0.203]	[0.207]	[0.210]	[0.200]	[0.193]
31-40 years old	0.0955	0.098	0.0783	0.0984	0.101	0.0499	0.059	0.601	0.607	0.524	0.616	0.517	0.437	0.202
	[0.0619]	[0.0623]	[0.0613]	[0.0616]	[0.0626]	[0.0611]	[0.0618]	[0.237]**	[0.238]**	[0.234]**	[0.236]***	[0.238]**	[0.232]*	[0.224]
41 and older	0.0488	0.0473	0.0828	0.052	0.13	0.0233	0.14	-0.00552	-0.0187	0.126	0.0429	-0.1	-0.077	-0.072
	[0.0644]	[0.0646]	[0.0641]	[0.0643]	[0.0843]	[0.0637]	[0.0836]*	[0.255]	[0.256]	[0.253]	[0.256]	[0.351]	[0.248]	[0.335]
Education														
Base: primary school or less														
Some highsch. (7-11 yrs)	0.00786	0.00718	-0.0002	0.00587	0.0201	-0.00108	0.00828	-0.169	-0.174	-0.201	-0.149	-0.107	-0.184	-0.127
	[0.0414]	[0.0416]	[0.0408]	[0.0414]	[0.0413]	[0.0409]	[0.0404]	[0.159]	[0.160]	[0.156]	[0.159]	[0.159]	[0.155]	[0.148]
Highschool or more (12+)	-0.0872	-0.0914	-0.0932	-0.0844	-0.0653	-0.0867	-0.0642	-0.287	-0.303	-0.285	-0.241	-0.198	-0.301	-0.171
	[0.0535]	[0.0545]*	[0.0533]*	[0.0537]	[0.0534]	[0.0525]*	[0.0527]	[0.222]	[0.226]	[0.223]	[0.222]	[0.225]	[0.216]	[0.213]
Marital Status														
Married	0.0748	0.0759	0.0436	0.0744	0.0694	0.0625	0.0388	0.252	0.247	0.178	0.189	0.0303	0.247	-0.0657
	[0.0443]*	[0.0445]*	[0.0439]	[0.0444]*	[0.0698]	[0.0437]	[0.0679]	[0.177]	[0.178]	[0.175]	[0.178]	[0.276]	[0.174]	[0.253]
Household characteristics														
Renting dwelling	0.185	0.191	0.171	0.171	0.156	0.189	0.146	0.0218	0.0512	0.0165	-0.00627	-0.0426	0.0826	-0.0644
	[0.0556]***	[0.0570]***	[0.0561]***	[0.0557]***	[0.0556]***	[0.0543]***	[0.0553]***	[0.241]	[0.245]	[0.239]	[0.241]	[0.240]	[0.233]	[0.226]
Base: household size 2-4														
Household size = 1	0.192	0.149	0.14	0.179	0.17	0.164	0.0672	0.565	0.454	0.497	0.483	0.302	0.497	-0.0302
	[0.0465]***	[0.0525]***	[0.0515]***	[0.0468]***	[0.0535]***	[0.0465]***	[0.0580]	[0.176]***	[0.196]**	[0.192]**	[0.179]***	[0.202]	[0.174]***	[0.208]
Household size >=5	-0.115	-0.0871	-0.072	-0.114	-0.0984	-0.121	-0.0714	-0.352	-0.198	-0.141	-0.334	-0.285	-0.396	-0.0958
	[0.0551]**	[0.0591]	[0.0577]	[0.0550]**	[0.0554]*	[0.0543]**	[0.0585]	[0.236]	[0.259]	[0.253]	[0.235]	[0.239]	[0.229]*	[0.250]
First time in the US														
Base: arrived on or before 2002														
Arrived btw 02 & 05	0.067	0.0638	0.0657	0.065	0.0573	0.0495	0.0378	0.424	0.428	0.442	0.4	0.377	0.374	0.279
	[0.0453]	[0.0455]	[0.0445]	[0.0452]	[0.0460]	[0.0443]	[0.0445]	[0.185]**	[0.185]**	[0.181]**	[0.184]**	[0.186]**	[0.178]**	[0.171]
Arrived after 2005	0.171	0.168	0.154	0.168	0.148	0.145	0.103	0.669	0.675	0.676	0.638	0.524	0.611	0.378
	[0.0477]***	[0.0480]***	[0.0469]***	[0.0477]***	[0.0492]***	[0.0468]***	[0.0479]**	[0.197]***	[0.197]***	[0.193]***	[0.196]***	[0.204]**	[0.190]***	[0.190]**

	Sent transfers							Log amount sent						
	base	**hh res**	**ind res**	**pback**	**relatives**	**financial**	**all**	**base**	**hh res**	**ind res**	**pback**	**relatives**	**financial**	**all**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Resources														
Base: first quartile of log household pce														
Log pce - second quartile		-0.0359	-0.0318				4.19E-05		-0.363	-0.355				-0.385
		[0.0684]	[0.0668]				[0.0667]		[0.281]	[0.274]				[0.261]
Log pce - third quartile		-0.00365	0.00856				0.0108		-0.063	-0.0537				-0.208
		[0.0666]	[0.0652]				[0.0651]		[0.267]	[0.261]				[0.248]
Log pce - fourth quartile		0.0747	0.0528				0.0657		0.109	0.0365				0.0476
		[0.0642]	[0.0634]				[0.0641]		[0.242]	[0.235]				[0.227]
Base: first quartile of log individual labor income														
Log li - second quartile			0.0711				0.0512			0.436				0.553
			[0.0610]				[0.0604]			[0.236]*				[0.223]**
Log li - third quartile			0.108				0.0948			0.873				1.076
			[0.0600]*				[0.0597]			[0.233]***				[0.222]***
Log li - fourth quartile			0.109				0.098			0.772				1.113
			[0.0607]*				[0.0627]			[0.232]***				[0.227]***
Expectation of returning to live to Mexico														
Base: does not expect to return														
Expects to return [...]														
w/some prob				0.1			0.0965				0.222			0.42
				[0.0694]			[0.0681]				[0.293]			[0.273]
with certainty				0.129			0.121				0.561			0.661
				[0.0724]*			[0.0729]*				[0.298]*			[0.284]**

	Sent transfers							Log amount sent						
	base	**hh res**	**ind res**	**pback**	**relatives**	**financial**	**all**	**base**	**hh res**	**ind res**	**pback**	**relatives**	**financial**	**all**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Relatives in Mexico														
Has spouse in Mx					0.0663		0.11					0.548		0.602
					[0.0814]		[0.0831]					[0.309]*		[0.300]**
Has mother or father in Mx					0.21		0.182					0.525		0.336
					[0.0537]***		[0.0526]***					[0.223]**		[0.206]
Has children					0.00421		-0.0251					-0.0957		-0.309
					[0.0616]		[0.0603]					[0.255]		[0.239]
Has ch. 0-15 yrs old in Mx					-0.0328		-0.0247					0.238		0.511
					[0.0670]		[0.0656]					[0.268]		[0.251]**
Has children 16+ in Mx					-0.0566		-0.0578					0.334		0.476
					[0.0813]		[0.0795]					[0.327]		[0.312]
Financial indicators														
Base: has no savings														
Only has savings in Mx						0.172	0.155						0.82	0.79
						[0.0643]***	[0.0637]**						[0.227]***	[0.218]***
Only has savings in US						0.0722	0.0513						0.168	0.0862
						[0.0465]	[0.0464]						[0.179]	[0.172]
Has sav. in Mx and US						0.138	0.134						0.762	0.915
						[0.0822]*	[0.0812]*						[0.290]***	[0.275]***
Base: does not have assets														
Only has assets in Mx						0.122	0.0939						0.667	0.787
						[0.0781]	[0.0787]						[0.325]**	[0.310]**
Only has assets in US						0.0586	0.0528						0.363	0.559
						[0.0597]	[0.0613]						[0.274]	[0.274]**
Has assets in Mx and US						0.181	0.146						0.62	0.756
						[0.0640]***	[0.0647]**						[0.276]**	[0.273]***

	Sent transfers							Log amount sent						
	base	**hh res**	**ind res**	**pback**	**relatives**	**financial**	**all**	**base**	**hh res**	**ind res**	**pback**	**relatives**	**financial**	**all**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Joint tests														
p-value age	0.00	0.00	0.01	0.00	0.00	0.00	0.08	0.06	0.04	0.18	0.06	0.17	0.14	0.84
p-value education	0.14	0.13	0.13	0.17	0.22	0.17	0.31	0.37	0.35	0.32	0.49	0.65	0.31	0.62
p-value household size	0.00	0.00	0.00	0.00	0.00	0.00	0.18	0.00	0.03	0.02	0.00	0.09	0.00	0.93
p-value location of int.	0.07	0.07	0.10	0.05	0.05	0.04	0.03	0.40	0.50	0.75	0.36	0.32	0.45	0.63
p-value date of interview	0.10	0.21	0.13	0.11	0.17	0.03	0.02	0.27	0.42	0.37	0.36	0.60	0.53	0.75
p-value arrival year	0.00	0.00	0.00	0.00	0.01	0.01	0.09	0.00	0.00	0.00	0.01	0.03	0.01	0.12
p-value savings						0.03	0.065						0.00	0.00
p-value assets						0.01	0.08						0.08	0.03
p-value household pce		0.34	0.63				0.66		0.36	0.488				0.27
p-value relatives in Mx					0.00		0.01					0.01		0.00
p-value children in Mx					0.67		0.70					0.364		0.03
p-value exp. to return				0.20			0.25				0.04			0.05
p-value labor income			0.26				0.36			0.00				0.00
Observations	529	529	529	529	529	529	529	371	371	371	371	371	371	371
R-squared	0.252	0.257	0.299	0.264	0.284	0.299	0.368	0.165	0.173	0.229	0.18	0.204	0.244	0.38

Notes: Standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1

Source: MxFLS3

Table 41: Transfers sent to Mexico in the last year by female migrants

	Sent transfers							Log amount sent						
	base (1)	**hh res (2)**	**ind res (3)**	**pback (4)**	**relatives (5)**	**financial (6)**	**all (7)**	**base (1)**	**hh res (2)**	**ind res (3)**	**pback (4)**	**relatives (5)**	**financial (6)**	**all (7)**
Age														
Base: 26-30 years old														
15-20 years old	-0.254	-0.267	-0.2	-0.261	-0.147	-0.257	-0.112	-0.738	-0.788	-0.64	-0.714	-0.882	-0.703	-0.741
	[0.0919]***	[0.0922]***	[0.0922]**	[0.0928]***	[0.0923]	[0.0910]***	[0.0944]	[0.397]*	[0.399]*	[0.399]	[0.404]*	[0.394]**	[0.393]*	[0.406]*
21-25 years old	0.0153	0.0041	-0.0101	0.0102	0.0945	0.00604	0.053	-0.52	-0.495	-0.552	-0.511	-0.46	-0.552	-0.513
	[0.0760]	[0.0768]	[0.0751]	[0.0774]	[0.0736]	[0.0754]	[0.0748]	[0.265]*	[0.273]*	[0.269]**	[0.273]*	[0.255]*	[0.266]**	[0.268]*
31-40 years old	0.00581	-0.00928	-0.0279	0.00282	0.00543	-0.0222	-0.0414	0.26	0.217	0.165	0.258	0.0904	0.141	-0.0929
	[0.0734]	[0.0738]	[0.0721]	[0.0738]	[0.0705]	[0.0735]	[0.0707]	[0.266]	[0.271]	[0.267]	[0.270]	[0.256]	[0.267]	[0.265]
41 and older	0.0101	-0.0164	-0.0319	0.0101	0.111	0.00853	0.0646	0.423	0.392	0.239	0.442	0.408	0.185	0.104
	[0.0874]	[0.0886]	[0.0866]	[0.0876]	[0.0978]	[0.0915]	[0.102]	[0.316]	[0.327]	[0.321]	[0.321]	[0.353]	[0.333]	[0.379]
Education														
Base: primary school or less														
Some highsch. (7-11 yrs)	0.0979	0.0683	0.0387	0.111	0.0896	0.07	0.0452	0.279	0.244	0.125	0.247	0.385	0.217	0.166
	[0.0618]	[0.0636]	[0.0627]	[0.0627]*	[0.0591]	[0.0627]	[0.0619]	[0.221]	[0.231]	[0.229]	[0.226]	[0.211]*	[0.225]	[0.229]
Highschool or more (12+)	0.0174	-0.0148	-0.0603	0.027	0.0251	0.00524	-0.0314	0.0846	0.0381	-0.166	0.0623	0.103	0.112	-0.124
	[0.0719]	[0.0735]	[0.0725]	[0.0725]	[0.0684]	[0.0720]	[0.0697]	[0.274]	[0.278]	[0.278]	[0.281]	[0.261]	[0.277]	[0.271]
Marital Status														
Married	0.197	0.179	0.24	0.172	0.0898	0.187	0.148	0.487	0.468	0.71	0.472	0.226	0.603	0.689
	[0.0667]***	[0.0674]***	[0.0671]***	[0.0681]**	[0.0766]	[0.0668]***	[0.0788]*	[0.287]*	[0.288]	[0.289]**	[0.292]	[0.304]	[0.289]**	[0.318]**
Household characteristics														
Renting dwelling	0.04	0.0521	0.0168	0.0219	0.0233	0.0258	-0.024	0.221	0.253	0.0981	0.247	0.0933	0.178	-6.4E-05
	[0.0661]	[0.0667]	[0.0653]	[0.0665]	[0.0630]	[0.0656]	[0.0627]	[0.238]	[0.241]	[0.239]	[0.241]	[0.229]	[0.236]	[0.231]
Base: household size 2-4														
Household size = 1	0.211	0.183	0.134	0.19	0.0535	0.192	-0.00262	1.662	1.548	1.474	1.612	0.916	1.732	0.978
	[0.115]*	[0.119]	[0.117]	[0.115]	[0.117]	[0.114]*	[0.116]	[0.409]***	[0.418]***	[0.413]***	[0.416]***	[0.412]**	[0.415]***	[0.416]**
Household size >=5	-0.124	-0.0843	-0.0529	-0.131	-0.092	-0.108	-0.0478	-0.521	-0.435	-0.42	-0.526	-0.268	-0.505	-0.243
	[0.0619]**	[0.0656]	[0.0650]	[0.0619]**	[0.0596]	[0.0618]*	[0.0628]	[0.232]**	[0.245]*	[0.242]*	[0.234]**	[0.228]	[0.234]**	[0.242]

	Sent transfers							Log amount sent						
	base	hh res	ind res	pback	relatives	financial	all	base	hh res	ind res	pback	relatives	financial	all
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(1)	(2)	(3)	(4)	(5)	(6)	(7)
First time in the US														
Base: arrived on or before 2002														
Arrived btw 02 & 05	0.0281	0.0396	0.0438	0.0271	-0.0118	0.0213	0.00198	-0.05	-0.0662	0.0023	-0.0499	0.0087	-0.0519	0.0104
	[0.0583]	[0.0582]	[0.0563]	[0.0580]	[0.0556]	[0.0573]	[0.0536]	[0.221]	[0.226]	[0.222]	[0.222]	[0.211]	[0.221]	[0.212]
Arrived after 2005	0.0741	0.0683	0.0745	0.0566	0.013	0.0649	0.000294	0.253	0.18	0.257	0.253	0.252	0.183	0.195
	[0.0601]	[0.0602]	[0.0581]	[0.0602]	[0.0581]	[0.0595]	[0.0564]	[0.227]	[0.233]	[0.229]	[0.227]	[0.216]	[0.225]	[0.218]
Resources														
Base: first quartile of log household pce														
Log pce - second quartile		0.0898	0.0533				0.0378		0.246	0.168				0.178
		[0.0674]	[0.0655]				[0.0626]		[0.264]	[0.260]				[0.255]
Log pce - third quartile		0.176	0.111				0.0785		0.105	-0.0506				0.0729
		[0.0697]**	[0.0680]				[0.0659]		[0.267]	[0.264]				[0.263]
Log pce - fourth quartile		0.203	0.128				0.0946		0.417	0.213				0.181
		[0.0755]***	[0.0738]*				[0.0722]		[0.278]	[0.277]				[0.273]
Base: first quartile of log individual labor income														
Log li - second quartile			0.0535				0.0148			0.136				0.00548
			[0.0718]				[0.0685]			[0.252]				[0.242]
Log li - third quartile			0.133				0.153			0.212				0.179
			[0.0865]				[0.0834]*			[0.288]				[0.284]
Log li - fourth quartile			0.277				0.287			0.762				0.806
			[0.106]***				[0.104]***			[0.336]**				[0.342]**
Expectation of returning to live to Mexico														
Base: does not expect to return														
Expects to return [...]														
w/some prob				0.154			0.0885				-0.366			-0.38
				[0.0728]**			[0.0665]				[0.299]			[0.282]
with certainty				0.166			0.0806				-0.322			-0.389
				[0.0809]**			[0.0747]				[0.316]			[0.307]

	Sent transfers							Log amount sent						
	base	**hh res**	**ind res**	**pback**	**relatives**	**financial**	**all**	**base**	**hh res**	**ind res**	**pback**	**relatives**	**financial**	**all**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Relatives in Mexico														
Has spouse in Mx														
Has mother or father in Mx					0.371		0.341					0.967		0.996
					[0.0643]***		[0.0645]***					[0.281]***		[0.287]***
Has children					0.0239		0.0199					-0.457		-0.432
					[0.0743]		[0.0767]					[0.277]		[0.299]
Has ch. 0-15 yrs old in Mx					0.298		0.282					1.329		1.336
					[0.104]***		[0.103]***					[0.321]***		[0.325]***
Has children 16+ in Mx					0.0355		0.047					0.782		0.736
					[0.112]		[0.112]					[0.410]*		[0.430]*
Financial indicators														
Base: has no savings														
Only has savings in Mx						0.117	-0.0291						0.72	0.312
						[0.125]	[0.123]						[0.410]*	[0.417]
Only has savings in US						0.0516	0.0479						0.213	0.152
						[0.0688]	[0.0664]						[0.246]	[0.240]
Has sav. in Mx and US						0.105	-0.145						0.511	-0.236
						[0.162]	[0.161]						[0.490]	[0.497]
Base: does not have assets														
Only has assets in Mx						-0.275	-0.217						-0.506	-0.343
						[0.152]*	[0.146]						[0.838]	[0.820]
Only has assets in US						0.0297	-0.0547						0.0224	-0.0283
						[0.0772]	[0.0740]						[0.341]	[0.331]
Has assets in Mx and US						0.161	0.0915						0.462	0.499

	Sent transfers							Log amount sent						
	base	hh res	ind res	pback	relatives	financial	all	base	hh res	ind res	pback	relatives	financial	all
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Joint tests														
p-value age	0.03	0.03	0.24	0.02	0.06	0.03	0.30	0.00	0.01	0.02	0.00	0.02	0.04	0.19
p-value education	0.22	0.36	0.30	0.16	0.28	0.44	0.44	0.41	0.50	0.48	0.50	0.15	0.62	0.44
p-value household size	0.01	0.10	0.33	0.01	0.23	0.03	0.75	0.00	0.00	0.00	0.00	0.04	0.00	0.04
p-value location of int.	0.50	0.56	0.65	0.44	0.45	0.45	0.57	0.14	0.17	0.20	0.15	0.12	0.18	0.18
p-value date of interview	0.06	0.09	0.17	0.06	0.15	0.29	0.31	0.13	0.26	0.47	0.14	0.22	0.43	0.61
p-value arrival year	0.78	0.93	0.86	0.91	0.82	0.68	0.95	0.33	0.54	0.38	0.33	0.34	0.49	0.46
p-value savings						0.69	0.67						0.28	0.75
p-value assets						0.02	0.04						0.23	0.12
p-value household pce		0.29	0.79				0.93		0.40	0.65				0.80
p-value relatives in Mx					0.00		0.00					0.00		0.00
p-value children in Mx					0.02		0.02					0.00		0.00
p-value exp. to return				0.12			0.46				0.79			0.57
p-value labor income			2.59	0	0	0	3.34			2.65	0	0	0	3.53
Observations	462	462	462	462	462	462	462	256	256	256	256	256	256	256
R-squared	0.161	0.18	0.247	0.175	0.254	0.213	0.36	0.201	0.213	0.263	0.214	0.307	0.252	0.405

Notes: Standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1

Source: MxFLS3

Table 42: Transfers sent to Mexico in the last year - Interaction between each covariate and a female dummy

	Sent transfers							Log amount sent						
	base	hh res	ind res	pback	relatives	financial	all	base	hh res	ind res	pback	relatives	financial	all
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Age														
15-20 years old	0.0351	0.0232	-0.0304	0.000917	0.161	0	0.0458	-2.061	-1.68	-1.821	-1.776	-1.847	-3.576	-4.251
	[0.250]	[0.251]	[0.246]	[0.253]	[0.260]	0	[0.144]	[1.607]	[1.592]	[1.568]	[1.587]	[1.576]	[1.909]*	[1.804]**
21-25 years old	0.115	0.104	0.0289	0.0796	0.27	0.0877	0.124	-2.006	-1.561	-1.794	-1.747	-1.576	-3.697	-4.228
	[0.246]	[0.246]	[0.240]	[0.249]	[0.255]	[0.0985]	[0.128]	[1.578]	[1.564]	[1.536]	[1.560]	[1.555]	[1.888]*	[1.788]**
26-30 years old	0.101	0.0996	0.0312	0.0719	0.184	0.0843	0.0832	-1.401	-1.003	-1.209	-1.118	-0.977	-3.082	-3.608
	[0.244]	[0.244]	[0.238]	[0.247]	[0.255]	[0.107]	[0.130]	[1.572]	[1.555]	[1.530]	[1.550]	[1.552]	[1.883]	[1.786]**
31-40 years old	-0.0183	-0.0356	-0.1	-0.0546	0.0678	-0.0181	-0.0404	-1.699	-1.35	-1.506	-1.424	-1.371	-3.35	-3.862
	[0.245]	[0.246]	[0.240]	[0.248]	[0.252]	[0.110]	[0.121]	[1.577]	[1.567]	[1.534]	[1.562]	[1.552]	[1.875]*	[1.775]**
41 and older	0.0344	0.0039	-0.108	0.00154	0.164	0.0437	0	-0.979	-0.597	-1.088	-0.74	-0.45	-2.84	-3.435
	[0.252]	[0.253]	[0.248]	[0.255]	[0.265]	[0.125]	0	[1.562]	[1.562]	[1.520]	[1.555]	[1.579]	[1.837]	[1.785]*
Education														
Primary or less (0-6 yrs)	-0.0959	-0.0575	-0.0119	-0.0977	0	0	0	0	-0.326	0	-0.289	0	0	0
	[0.0867]	[0.0884]	[0.0868]	[0.0872]	0	0	0	0	[0.350]	0	[0.348]	0	0	0
Some highsch. (7-11 yrs)	-0.0114	-0.00901	0.0141	-0.00174	0.0781	0.0622	0.0393	0.386	0.0319	0.281	0.0436	0.409	0.354	0.274
	[0.0795]	[0.0799]	[0.0781]	[0.0794]	[0.0697]	[0.0715]	[0.0707]	[0.268]	[0.316]	[0.273]	[0.313]	[0.263]	[0.266]	[0.263]
Highschool or more (12+)	0	0	0	0	0.0917	0.0846	0.027	0.357	0	0.149	0	0.296	0.413	0.132
	0	0	0	0	[0.0846]	[0.0859]	[0.0847]	[0.343]	0	[0.346]	0	[0.340]	[0.339]	[0.333]
Marital Status														
Married	0.128	0.106	0.202	0.105	0.0136	0.128	0.0768	0.336	0.325	0.644	0.381	0.523	0.454	0.988
	[0.0774]*	[0.0778]	[0.0771]***	[0.0782]	[0.0946]	[0.0768]*	[0.0940]	[0.331]	[0.333]	[0.331]*	[0.336]	[0.385]	[0.327]	[0.372]***
Household characteristics														
Renting dwelling	-0.128	-0.119	-0.145	-0.138	-0.128	-0.148	-0.187	0.327	0.327	0.194	0.389	0.245	0.226	0.154
	[0.0836]	[0.0850]	[0.0834]*	[0.0839]*	[0.0819]	[0.0821]*	[0.0815]**	[0.334]	[0.339]	[0.333]	[0.336]	[0.329]	[0.326]	[0.319]
Household size = 1	0.102	0.4	0.442	0.383	-0.206	0.108	0.535	0.811	0.458	1.084	-0.289	0.253	1.072	0.34
	[0.134]	[0.416]	[0.405]	[0.227]*	[0.164]	[0.147]	[0.466]	[0.554]	[1.415]	[1.381]	[1.066]	[0.684]	[0.623]*	[1.699]
Household size 2-4	0.0516	0.337	0.431	0.339	-0.103	0.0597	0.612	-0.448	-0.802	-0.0331	-1.609	-0.636	-0.298	-0.845
	[0.0968]	[0.406]	[0.395]	[0.207]	[0.119]	[0.115]	[0.452]	[0.445]	[1.377]	[1.344]	[0.988]	[0.550]	[0.545]	[1.657]
Household size >=5	0.0602	0.37	0.483	0.344	-0.0716	0.0939	0.676	-0.647	-1.068	-0.303	-1.844	-0.652	-0.423	-0.95
	[0.110]	[0.412]	[0.401]	[0.215]	[0.126]	[0.124]	[0.458]	[0.499]	[1.409]	[1.374]	[1.028]*	[0.589]	[0.575]	[1.697]

	Sent transfers							Log amount sent						
	base	**hh res**	**ind res**	**pback**	**relatives**	**financial**	**all**	**base**	**hh res**	**ind res**	**pback**	**relatives**	**financial**	**all**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
First time in the US														
Arrived on or before 02	-0.0629	-0.0612	-0.0189	0.00562	-0.177	-0.0325	-0.0852	0.587	0.582	0.529	0.556	0.209	0.951	0.623
	[0.225]	[0.225]	[0.220]	[0.228]	[0.231]	[0.221]	[0.225]	[1.536]	[1.538]	[1.494]	[1.533]	[1.505]	[1.494]	[1.404]
Arrived btw 02 & 05	-0.166	-0.141	-0.0971	-0.0906	-0.268	-0.117	-0.132	0.167	0.148	0.131	0.147	-0.0997	0.573	0.386
	[0.225]	[0.225]	[0.220]	[0.228]	[0.231]	[0.222]	[0.226]	[1.529]	[1.531]	[1.488]	[1.527]	[1.498]	[1.487]	[1.397]
Arrived after 05	-0.231	-0.227	-0.155	-0.173	-0.335	-0.164	-0.185	0.13	0.0407	0.067	0.121	-0.118	0.513	0.421
	[0.225]	[0.225]	[0.220]	[0.228]	[0.231]	[0.222]	[0.225]	[1.533]	[1.535]	[1.492]	[1.531]	[1.502]	[1.490]	[1.402]
Resources														
Log pce - first quartile		-0.404	-0.367				-0.223		0.061	-0.106				0.12
		[0.405]	[0.397]				[0.389]		[1.350]	[1.335]				[1.289]
Log pce - second quartile		-0.262	-0.263				-0.167		0.697	0.427				0.677
		[0.404]	[0.397]				[0.389]		[1.356]	[1.342]				[1.290]
Log pce - third quartile		-0.219	-0.274				-0.158		0.324	-0.0336				0.508
		[0.403]	[0.395]				[0.387]		[1.359]	[1.346]				[1.292]
Log pce - fourth quartile		-0.297	-0.313				-0.231		0.438	0.144				0.321
		[0.399]	[0.392]				[0.384]		[1.339]	[1.332]				[1.271]
Log li - first quartile			-0.0644				-0.0105			-0.0421				0.327
			[0.0952]				[0.0942]			[0.440]				[0.424]
Log li - second quartile			-0.0809				-0.0689			-0.371				-0.306
			[0.106]				[0.105]			[0.469]				[0.451]
Log li - third quartile			0.0105				0.0819			-0.656				-0.458
			[0.116]				[0.115]			[0.481]				[0.464]
Log li - fourth quartile			0.157				0.213			0.209				0.314
			[0.131]				[0.134]			[0.520]				[0.513]
Expectation of returning to live to Mexico														
w/zero probability				-0.377			-0.449				1.53			0.959
				[0.212]*			[0.220]**				[0.994]			[0.987]
w/some probability				-0.283			-0.424				1.298			0.303
				[0.195]			[0.206]**				[0.919]			[0.921]
w/certainty				-0.307			-0.458				0.912			-0.0415
				[0.199]			[0.211]**				[0.926]			[0.929]

	Sent transfers							Log amount sent						
	base	**hh res**	**ind res**	**pback**	**relatives**	**financial**	**all**	**base**	**hh res**	**ind res**	**pback**	**relatives**	**financial**	**all**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Relatives in Mexico														
Has spouse in Mx														
Has mother or father in Mx					0.167		0.162					0.383		0.579
					[0.0805]**		[0.0797]**					[0.351]		[0.339]*
Has children					0.000733		0.0459					-0.382		-0.0575
					[0.0941]		[0.0943]					[0.373]		[0.371]
Has ch. 0-15 yrs old in Mx					0.352		0.312					1.144		0.815
					[0.120]***		[0.118]***					[0.419]***		[0.404]**
Has children 16+ in Mx					0.0732		0.0952					0.44		0.201
					[0.136]		[0.134]					[0.527]		[0.520]
Financial indicators														
Has no savings						-0.487	-0.218						0.173	1.177
						[0.300]	[0.299]						[0.557]	[0.559]**
Only has savings in Mx						-0.505	-0.354						0.0887	0.786
						[0.323]	[0.321]						[0.652]	[0.650]
Only has savings in US						-0.489	-0.21						0.286	1.291
						[0.311]	[0.310]						[0.591]	[0.592]**
Has sav. in Mx and US														
Has no assets						-0.497	-0.485						0	0
						[0.349]	[0.338]						0	0
Only has assets in Mx						0.437	0.327						1.388	1.167
						[0.162]***	[0.158]**						[0.864]	[0.840]
Only has assets in US						0.409	0.235						0.967	0.509
						[0.158]***	[0.157]						[0.830]	[0.818]
Has assets in Mx and US						0.437	0.294						1.248	0.878
						[0.156]***	[0.155]*						[0.825]	[0.809]

	Sent transfers							Log amount sent						
	base	**hh res**	**ind res**	**pback**	**relatives**	**financial**	**all**	**base**	**hh res**	**ind res**	**pback**	**relatives**	**financial**	**all**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Joint tests														
p-value age	0.73	0.65	0.59	0.71	0.36	0.71	0.45	0.09	0.12	0.31	0.11	0.10	0.13	0.08
p-value education	0.42	0.75	0.93	0.36	0.45	0.56	0.86	0.33	0.42	0.59	0.46	0.30	0.34	0.57
p-value household size	0.89	0.77	0.66	0.39	0.64	0.86	0.40	0.02	0.03	0.06	0.00	0.17	0.01	0.06
p-value arrival year	0.18	0.20	0.37	0.17	0.16	0.40	0.58	0.46	0.37	0.47	0.50	0.72	0.48	0.84
p-value savings						0.61	0.41						0.95	0.15
p-value assets						0.04	0.16						0.30	0.22
p-value household pce					0.00		0.01					0.03		0.08
p-value relatives in Mx					0.01		0.02					0.02		0.12
p-value children in Mx				0.35			0.18				0.23			0.16
p-value exp. to return		0.37	0.74				0.87		0.61	0.68				0.63
p-value labor income			0.38				0.23			0.29				0.16
Observations	895	895	895	895	895	895	895	585	585	585	585	585	585	585
R-squared	0.21	0.22	0.27	0.23	0.27	0.26	0.36	0.27	0.28	0.33	0.28	0.33	0.33	0.46

Notes: Standard errors in brackets. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Source: MxFLS3

References

Adato, M., Coady, D., and Ruel, M. (2000), "An Operations Evaluation of PROGRESA from the Perspective of Beneficiaries, Promotoras, School Directors, and Health Staff," Final Report, International Food Policy Research Institute, Washington D.C.

Albarran, P. and Attanasio, O.P. (2005), "Do Public Transfers Crowd Out Private Transfers? Evidence from a Randomized Experiment in Mexico," in *Insurance Against Poverty*, ed. S. Dercon, pp. 281-304, Oxford University Press.

Altonji, J. G., Hayashi, F., and Kotlikoff, L. J. (1992), "Is the Extended Family Altruistically Linked? Direct Tests Using Micro Data," *American Economic Review*, 82, 1177-1198.

Altonji, J. G., Hayashi, F., and Kotlikoff, L. J. (1997), "Parental Altruism and Inter Vivos Transfers: Theory and Evidence," *Journal of Political Economy*, 105, 1121-1166.

Amuedo-Dorantes, C. and Pozo, S. (2006), "Remittances as Insurance: Evidence from Mexican Immigrants," *Journal of Population Economics*, 19, 227-254.

Amuedo-Dorantes, C., Bansak, C., and Pozo, S. (2005), "On the Remitting Patterns of Immigrants: Evidence from Mexican Survey Data," *Economic Review*, 90, 37-58.

Angelucci, M. and Attanasio, O. (2006), "Estimating ATT Effects with Non-Experimental Data and Low Compliance," IZA Discussion Paper No. 2368.

Angelucci, M. and Attanasio, O. (2009), "Oportunidades: Program Effect on Consumption, Low Participation and Methodological Issues," *Economic Development and Cultural Change*, 57, 479-506.

Angelucci, M. and DeGiorgi, D. (2009), "Indirect Effects of an Aid Program: How Do Cash Transfers Affect Ineligibles' Consumption?," *American Economic Review*, 99, 486-508.

Angelucci, M., DeGiorgi, G., Rangel, M., and Rasul, I. (2009), "Insurance and Investment within Family Networks," Bread Working Paper No. 260.

Angelucci, M., DeGiorgi, G., Rangel, M., and Rasul, I. (2010), "Family Networks and School Enrollment: Evidence from a Randomized Social Experiment," *Journal of Public Economics*, 94, 197-221.

Ashraf, N. (2009), "Spousal Control and Intra-Household Decision Making: An Experimental Study in the Philipines," *American Economic Review*, 99, 1245-1277.

Ashraf, N., Field, E., and Lee, J. (2012), "Household Bargaining and Excess Fertility: An Experimental Study in Zambia," Working Paper.

Behrman, J. R. and Hoddinott, J. (2005), "Programme Evaluation with Unobserved Heterogeneity and Selective Implementation: The Mexican PROGRESA Impact on Child Nutrition," *Oxford Bulletin of Economics and Statistics*, 67, 547-569.

Behrman, J. R. and Todd, P. E. (1999), "Randomness in the Experimental Samples of PROGRESA (Education, Health, and Nutrition Program)," International Food Policy Research Institute, Washington D.C.

Bianchi, S. M., Hotz, V. J., McGarry, K., and Seltzer, J. A. (2007), "Intergenerational Ties: Alternative Theories, Empirical Findings and Trends, and Remaining Challenges," in *Intergenerational Caregiving*, eds. A. Booth, N. Crouter, S. Bianchi, and J. Seltzer, pp. 3-44, Urban Institute Press.

Bobonis, G. J. and Finan, F. (2006), "Endogenous peer effects in school participation," Working Paper.

Bourguignon, F., Browning, M., and Chiappori, P.-A. (2009), "Efficient Intra-Household Allocations and Distribution Factors: Implications and Identification," *Review of Economic Studies*, 76, 503-528.

Browning, M. and Chiappori, P.-A. (1998), "Efficient Intra-Household Allocations: A General Characterization and Empirical Tests," *Econometrica*, 66, 1241-1278.

Browning, M., Bourguignon, F., Chiappori, P.-A., and Lechene, V. (1994), "Incomes and Outcomes: A Structural Model of Intra-Household Allocation," *Journal of Political Economy*, 102, 1067-1096.

Browning, M., Chiapori, P.-A., and Weiss, Y. (2011), "Family Economics," Mimeo.

Castilla, C. (2010), "Show me the Money: Intra-household Allocation under Incomplete Information," Working Paper.

Castilla, C. and Walker, T. (2013), "Is Ignorance Bliss? The Effect of Asymmetric Information between Spouses on Intra-Household Allocations," *American Economic Review*, 103, 263-268.

Chen, J. (2006), "Migration and Imperfect Monitoring: Implications for Intra-Household Allocation," *American Economic Review*, 96, 227-231.

Chen, J. (2013), "Identifying Non-cooperative Behavior among Spouses: Child Outcomes in Migrant-Sending Households," *Journal of Development Economics*, 100, 1-18.

Chiappori, P.-A. (1988), "Rational Household Labor Supply," *Econometrica*, 56, 63-89.

Chiappori, P.-A. (1992), "Collective Labor Supply and Welfare," *Journal of Political Economy*, 100, 437-467.

Chiappori, P.-A., Samphantharak, K., Schulhofer-Wohl, S., and Townsend, R. M. (2011), "Heterogeneity and Risk Sharing in Thai Villages," NBER Working Paper No. 16696.

Chiquiar, D. and Hanson, G. H. (2005), "International migration, self-selection, and the distribution of wages: evidence from Mexico and the United states," *Journal of Political Economy*, 113, 239-281.

Cox, D. (1987), "Motives for Private Income Transfers," *Journal of Political Economy*, 95, 508-546.

Cox, D., Hansen, B. E., and Jimenez, E. (2004), "How Responsive are Private Transfers to Income? Evidence from a Laissez-Faire Economy," *Journal of Public Economics*, 88, 2193-2219.

Dalton, M., Hotz, V. J., and Thomas, D. (2012), "The Family that Shares is the Family that Cares: Are Extended Families Efficient?," Working Paper.

De la Briere, B., Sadoulet, E., de Janvry, A., and Lambert, S. (2002), "The Roles of Destination, Gender, and Household Composition in Explaining Remittances: An Analysis of the Dominican Sierra," *Journal of Development Economics*, 68, 309-328.

Duflo, E. (2001), "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment," *American Economic Review*, 91, 795-813.

Farfan, G., Genoni, M., Rubalcava, L., Teruel, G., Thomas, D., and Velasquez, A. (2013), "Mexicans in America," Working Paper.

Fernald, L. C. H., Gertler, P. J., and Neufeld, L. M. (2009), "10-year effect of Oportunidades, Mexico's Conditional Cash Transfer Programme, on Child Growth, Cognition, Language, and Behaviour: a Longitudinal Follow-Up Study," *The Lancet*, 374, 1997-2005.

Frankenberg, E., Suriastini, W., and Thomas, D. (2005), "Can Expanding Access to Basic Healthcare Improve Children's Health Status? Lessons from Indonesia's 'Midwife in the Village' Programme," *Population Studies*, 59, 5-19.

Gertler, P. J. (2004), "Do Conditional Cash Transfers Improve Child Health? Evidence from PROGRESA's Control Randomized Experiment," *American Economic Review*, 94, 336-341.

Gutiérrez, J. P., Bertozzi, S., and Gertler, P. J. (2003), "Evaluación de la Identificación de Familias Beneficiarias en el Medio Urbano," Evaluación de Resultados de Impacto del Programa de Desarrollo Humano Oportunidades, Instituto Nacional de Salud Pública, Mexico.

Hamoudi, A. (2010), "Risk Aversion, Household Partition, and Consumption Smoothing in Rural Mexico," Working Paper.

Hanson, G. H. (2006), "Illegal Migration from Mexico to the United States," *Journal of Economic Literature*, 44, 72-96.

Hanson, G. H. and McIntosh, C. (2008), "The Great Mexican Emigration," Working Paper.

Hayashi, F., Altonji, J. G., and Kotlikoff, L. J. (1996), "Risk-Sharing Between and Within Families," *Econometrica*, 64, 261-294.

Kaestner, R. and Malamud, O. (2014), "Self-Selection and International Migration: New Evidence from Mexico," *Review of Economics and Statistics*, 96, 78-91.

Kaplan, R. and Saccuzzo, D. (1997), *Psychological testing: Principles, Applications, and Issues*, Brooks/Cole, Pacific Grove, CA, 4th edn.

Kinnan, C. and Townsend, R. (2012), "Kinship and Financial Networks, Formal Financial Access, and Risk Reduction," *American Economic Review*, 102, 289-293.

LaFave, D. and Thomas, D. (2013), "Extended Families and Child Well-Being," Working Paper.

Leroy, J., García-Guerra, A., García, R., Dominguez, C., Rivera, J., and Neufeld, L. M. (2008), "The Oportunidades Program Increases the Linear Growth of Children Enrolled at Young Ages in Urban Mexico," *Journal of Nutrition*, 138, 793-798.

Leroy, J., Ruel, M., and Verhofstadt, E. (2009), "The Impact of Conditional Cash Transfer Programmes on Child Nutrition: A Review of Evidence Using a Programme Theory Framework," *Journal of Development Effectiveness*, 1, 103-129.

Lucas, R.E.B. and Stark, O. (1985), "Motivations to Remit: Evidence from Botswana," *Journal of Political Economy*, 93, 901-918.

Martorell, R. (1999), "The Nature of Child Malnutrition and its Long-Term Implications," *Food and Nutrition Bulletin*, 20, 288-292.

Martorell, R. and Habicht, J.-P. (1986), "Growth in Early Childhood in Developing Countries," in *Human Growth: A Comprehensive Treaise*, eds. F. Falkner and J.M. Tanner, pp. 241-262, New York: Plenum Press.

Martorell, R., Behrman, J. R., Flores, R., and Stein, A. D. (2005), "Rationale for a Follow-Up Focusing on Economic Productivity," *Food and Nutrition Bulletin*, 26, 5-14.

Massey, D. S. and Basem, L. C. (1992), "Determinants of Savings, Remittances, and Spending Patterns among U.S. Migrants in Four Mexican Communities," *Sociological Inquiry*, 62, 185-207.

Massey, D. S., Goldring, L., and Durand, J. (1994), "Continuities in Transnational Migration: An Analysis of Nineteen Mexican Communities," *American Journal of Sociology*, 99, 1942-1533.

Mazzocco, M. (2007), "Household Intertemporal Behaviour: A Collective Characterization and a Test of Commitment," *Review of Economic Studies*, 74, 857-895.

McKenzie, D. (2003), "How do Households Cope with Aggregate Shocks? Evidence from the Mexican Peso Crisis," *World Development*, 31, 1179-1199.

McKenzie, D. and Rapoport, H. (2007), "Network Effects and the Dynamics of Migration and Inequality: Theory and Evidence from Mexico," *Journal of Development Economics*, 84, 1-24.

McKenzie, D. and Rapoport, H. (2010), "Self-Selection Patterns in Mexico-U.S. Migration: The Role of Migration Networks," *Review of Economics and Statistics*, 92, 811-821.

McKenzie, D., Gibson, J., and Stillman, S. (2013), "A Land of Milk and Honey with Streets Paved with Gold: Do Emigrants Have Over-Optimistic Expectations about Incomes Abroad?," *Journal of Development Economics*, 102, 116-127.

Munshi, K. and Rosenzweig, M. (2009), "Why Is Mobility in India so Low? Social Insurance, Inequality and Growth," NBER Working Paper No. 14850.

Neufeld, L. M., Sotres-Álvarez, D., García-Peregrino, A., García-Guerra, L., Tolentino-Mayo, L. F., and Rivera-Dommarco, J. (2004a), "Evaluación del Estado Nutricional y Adquisición de Lenguaje en Niños de Localidades Urbanas con y sin el Programa Oportunidades," Documento Técnico No. 7, Evaluación Externa de Impacto del Programa de Desarrollo Humano Oportunidades, Instituto Nacional de Salud Pública, México.

Neufeld, L. M., Sotres-Álvarez, D., Gertler, P. J., Tolentino Mayo, L., Jiménez Ruiz, J., Fernald, L. C. H., Villalpando, S., Shamah, T., and Rivera-Dommarco, J. (2004b), "Impact of Oportunidades on Child Growth and Nutritional Status in Rural Communities," External Evaluation of the Impact of the Human Development Program Oportunidades, Instituto Nacional de Salud Pública, México.

Neufeld, L. M., García-Guerra, A., Leroy, J., Flores López, M. L., Fernández Gaxiola, A. C., and Rivera-Dommarco, J. A. (2005), "Impacto del programa Oportunidades en nutrición y alimentación en zona urbanas de México," Evaluación Externa, Instituto Nacional de Salud Pública, México.

Özden, Ç. and Schiff, M. (eds.) (2006), *International Migration, Remittances, and the Brain Drain*, The International Bank for Reconstruction and Development/The World Bank, Washington D.C.

Parker, S. W., Rubalcava, L., and Teruel, G. (2008), "Evaluating Conditional Schooling and Health Programs," in *Handbook of Development Economics*, eds. T. P. Schultz and J. Strauss, vol. 4, pp. 3963-4035, North-Holland.

Parker, S. W., Todd, P. E., and Wolpin, K. I. (2006), "Within Family Treatment Effect Estimators: The Impact of Oportunidades on Schooling in Mexico," Working Paper.

Rapoport, H. and Docquier, F. (2006), "The Economic of Migrants' Remittances," in *Handbook of the Economics of Giving, Altruism and Reciprocity*, eds. S.-C. Kolm and J. M. Ythier, vol. 2, pp. 1135-1198, Elsevier.

Raven, J. (2000), "The Raven's Progressive Matrices: Change and Stability over Culture and Time," *Cognitive Psychology*, 41, 1-48.

Rivera, J. A., Rodríguez, G., Shamah, T., Rosado, J. L., Casanueva, E., Maulén, I., Toussaint, G., and García-Aranda, A. (2000), "Implementation, Monitoring and Evaluation of the Nutrition Component of the Mexican Social Programme (PROGRESA)," *Food and Nutrition Bulletin*, 21, 35-42.

Rivera, J. A., Sotres-Alvarez, D., Habicht, J.-P., Shamah, T., and Villapando, S. (2004), "Impact of the Mexican Program for Education, Health, and Nutrition (Progresa) on Rates of Growth and Anemia in Infants and Young Children. A Randomized Effectiveness Study," *Journal of the American Medical Association*, 291, 2563-2570.

Rosenzweig, M. R. and Stark, O. (1989), "Consumption Smoothing, Migration, and Marriage: Evidence from Rural India," *Journal of Political Economy*, 97, 905-926.

Rubalca, L. and Teruel, G. (2006), "User's Guide for the Mexican Family Life Survey: First Wave," http://www.ennvih-mxfls.org.

Rubalca, L. and Teruel, G. (2008), "User's Guide for the Mexican Family Life Survey: Second Wave," http://www.ennvih-mxfls.org.

Rubalcava, L., Teruel, G., and Thomas, D. (2009), "Investment, Time Preferences, and Public Transfers Paid to Women," *Economic Development and Cultural Change*, 57, 507-538.

Schultz, T. P. (2004), "School Subsidies for the Poor: Evaluating a Mexican Strategy for Reducing Poverty," *Journal of Development Economics*, 74, 199-250.

Skoufias, E. (2005), "PROGRESA and Its Impacts on the Welfare of Rural Households in Mexico," International Food Policy Research Institute Research Report No. 139.

Skoufias, E., Davis, B., and Behrman, J. R. (1999a), "An Evaluation of the Selection of Beneficiary Households in the Education, Health, and Nutrition Program (PROGRESA) of Mexico," Final Report, International Food Policy Research Institute, Washington D.C.

Skoufias, E., Davis, B., and Behrman, J. R. (1999b), "Targeting the Poor in Mexico: An Evaluation of the Selection of Households into Progresa," International Food Policy Research Institute, Washington D.C.

Stark, O. and Lucas, R.E.B. (1988), "Migration, Remittances and the Family," *Economic Development and Cultural Change*, 36, 465-481.

Strauss, J. and Thomas, D. (1995), "Human Resources: Empirical Modeling of Household and Family Decisions," in *Handbook of Development Economics*, eds. J. Behrman and T.N. Srinivasan, vol. 3, pp. 1883-2023, Elsevier.

Teruel, G. and Rubalcava, L. (2007), "Attrition in PROGRESA," Working Paper.

Thomas, D. and Frankenberg, E. (2007), "Household Responses to the Financial Crisis in Indonesia: Longitudinal Evidence on Poverty, Resources and Well-Being," in *Globalization and Poverty*, ed. A. Harrison, University of Chicago Press.

Townsend, R. M. (1994), "Risk and Insurance in Village India," *Econometrica*, 62, 539-591.

Winters, P., de Janvry, A., and Sadoulet, E. (2001), "Family and Community Networks in Mexico-U.S. Migration," *Journal of Human Resources*, 36, 159-184.

Witoelar, F. (2013), "Risk Sharing within the Extended Family: Evidence from the Indonesia Family Life Survey," *Economic Development and Cultural Change*, 62, 65-94.

Yang, D. (2008), "International Migration, Remittances and Household Investment: Evidence from Philippine Migrant's Exchange Rate Shocks," *Economic Journal*, 118, 591-630.

Yang, D. (2011), "Migrant Remittances," *Journal of Economic Perspectives*, 25, 129-152.

Yang, D. and Choi, H. (2007), "Are Remittances Insurance? Evidence from Rainfall Shocks in the Philippines," *World Bank Economic Review*, 21, 219-248.

Biography

María Gabriela Farfán Bertrán was born in Mendoza, Argentina, on March 22nd, 1983. She obtained her Bachelor's degree in Economics, *Summa Cum Laude*, from Universidad Nacional de Cuyo in 2006. She then enrolled in the Master's in Economics program at Universidad Nacional de La Plata.

Gabriela started her graduate studies at Duke University in the fall of 2008. She received a Master of Arts in Economics in 2009, and will be receiving her Ph.D. in Economics in May 2014. To pursue her graduate studies, Gabriela received financial awards from Duke University, the Fred and Barbara Sutherland Fellowship and the Feibusch Family Fellowship; as well as the Hewlett/IIE Dissertation Fellowship awarded by the William and Flora Hewlett Foundation.

Gabriela will join the Development Research Group of the World Bank in Washington D.C. in May 2014.